DIAGNOSING PREDISPOSITION TO FAT DEPOSITION AND ASSOCIATED CONDITIONS

Cross-Reference to Related Applications

[0001] This application claims benefit of U.S. patent application no. 60/392,361 filed 27 June 2002. The contents of that application is incorporated herein by reference.

Field of the Invention

[0002] The invention relates to genetic alterations in nucleic acids that are associated with fat deposition and diabetes.

Background

[0003] Individuals who are obsese have excess body fat compared to set standards. Obesity can be determined by several methods including body mass index (BMI) measurements, weight-for height charts, and body fat measurements determined by skinfold thickness and bioelectrical impedance. Obesity affects 58 million people across the United States, which represents approximately one-quarter to one-third of the adult population, and its prevalence is increasing to epidemic proportions in the United States and in other industrialized nations.

[0004] Recognized since 1985 as a chronic disease, obesity-related medical conditions contribute to approximately 300,000 deaths each year, second only to smoking as a cause of preventable death. (JAMA, *276:* 1907-1915 (1996)). Obesity has been established as a major risk factor for type II diabetes melitus, hypertension, cardiovascular disease and some cancers in both men and women (JAMA, *282:* 1523-1529 (1999)). Other comorbid conditions include sleep apnea, osteoarthritis, infertility, idiopathic intracranial hypertension, lower extremity venous stasis disease, gastro-esophageal reflux and urinary stress incontinence.

[0005] The total cost attributable to obesity amounted to $99.2 billion in 1995. Approximately $51.65 billion of those dollars were direct medical costs. The cost of obesity to U.S. business in 1994 was estimated to total $12.7 billion, and health-related economic costs of obesity to businesses in the United States is substantial, representing approximately 5% of total medical care costs. (American Journal of Health Promotion, *13 (2):* 120-127 (1998)). It was found that as BMI increases, so do the number of sick days, medical claims and health care costs and that the mean annual health care costs for the BMI "at risk" population was $2,274 versus $1,499 for the "not at risk" group.

[0006] An accumulation of adipose tissue on the trunk and around the waist, known as central fat, also confers an increased risk of type II diabetes and cardiovascular disease (Lundgre*n et al., Int. J.*

Obes., 13(4): 413-23 (1989); Ohlson *et al., Diabetes, 34(10):* 1055-8 (1985)). In addition, central obesity has been implicated in a condition known as the metabolic syndrome (or syndrome X), which is associated with increased risk of cardiovascular disease, vascular dementia, and diabetes. The metabolic syndrom is a descriptive term for the coexistence of all of the following or differing combinations of central fat, hypertension, glucose intolerance, dyslipidemia (elevated triglycerides and low HDL cholesterol), and impaired insulin stimulated glucose uptake ("insulin resistance"). Prevalence of central fat and its relationship to general obesity differs between ethnic groups and gender (McKeigue *et al., Diabetologia, 35(8):* 785-91 (1992); McKeigue *et al., Lancet, 337(8738)*: 382-6 (1991)). A majority of male subjects having high central fat are also obese in terms of BMI, and obese subjects often have a central distribution of fat, which suggests an overlap between these two conditions. While this relationship is not as strongly correlated in women, central fat increases after menopause.

[0007] Current anti-obesity therapeutics (*e.g.*, Phentermine, Sibutramine, and Orlistat) are largely ineffective and there is an urgent need to define the etiology of this disease and initiate rational, mechanism-based drug development. Mouse QTL and human studies have postulated that the 12q22 to q23 region, and specifically the insulin-like growth factor 1 (IGF-1) gene in that region, play a role in body weight regulation and viseral fat deposition (Collins, A.C. *et al., Mamm. Genome, 4:* 454-458 (1993); Sun, G. *et al., Int. J. Obes., 23:* 929-935 (1999); Keightley, P.D. *et al., Genetics, 142:* 227-235 (1996). Also, other studies have linked obesity with certain portions of the human genome (Perusse, L. *et al., Obesity Research, 9:* 135-169 (2001); Chagnon, Y.C. *et al., Obesity Research, 8:* 89-117 (2000)). Specifically, the CD36L gene on chromosome 12 was implicated in plasma lipid levels and with BMI (Acton, S. *et al., Arterioscler. Thromb. Vasc. Biol., 19:* 1734-1743 (1999)), the 12q24 chromosomal region was postulated as playing a role in obesity in a Quebec Family Study (Perusse, L. *et al., Diabetes, 50:* 614-621 (2001)), and it was reported that certain polymorphic loci on chromosome four are associated with obesity (Stone *et al., American J. Human Genetics*, "A major predisposition locus for severe obesity at 4p15-p14," June 2002).

Summary

[0008] It has been discovered that polymorphic variations in a gene encoding a phospholipase A2 polypeptide known as PLA2G1B, which is located on chromosome twelve, are associated with central fat deposition. Thus, featured herein is a method for diagnosing predisposition to fat deposition or leanness in a subject, which comprises detecting the presence or absence of one or more polymorphic variations in a PLA2G1B nucleotide sequence in a nucleic acid sample from a subject, where the PLA2G1B nucleotide sequence is set forth as SEQ ID NO: 1 or a substantially identical nucleotide sequence thereof. In embodiments of the invention, polymorphic variations at positions 7328 or 9182 in SEQ ID NO: 1 may be detected.

[0009] In addition, it was discovered that a polymorphic variation at position 7256 was associated with type II diabetes (non-insulin dependent diabetes mellitus, or NIDDM) in subjects. Therefore, featured herein is a method for diagnosing predisposition to NIDDM in a subject, which comprises obtaining a nucleic acid sample from the subject and detecting the presence or absence of a polymorphic variation in a PLA2 nucleotide sequence associated with NIDDM. Also featured are methods for identifying candidate therapeutic molecules for treating NIDDM in a subject and methods for treating NIDDM in a subject by administering such a therapeutic molecule.

[0010] Also featured herein are nucleic acids that encode a PLA2G1B polypeptide and include one or more polymorphic variations at positions 7256, 7328 and/or 9182 in SEQ ID NO: 1 associated with cetral fat deposition, leaness and/or NIDDM, and oligonucleotides which hybridize to those nucleic acids.

Brief Description Of The Drawings

[0011] Figures 1A-1D depict the PLA2G1B nucleotide sequence reported as SEQ ID NO:1. The following nucleotide representations are used throughout: "A" or "a" is adenosine, adenine, or adenylic acid; "C" or "c" is cytidine, cytosine, or cytidylic acid; "G" or "g" is guanosine, guanine, or guaylic acid; "T" or "t" is thymidine, thymine, or thymidylic acid; and "I" or "i" is inosine, hypoxanthine, or inosinic acid. Exons are indicated in italicized lower case type, introns are depicted in normal text lower case type, and polymorphic sites are depicted in bold upper case type. SNPs are designated by the following convention: "R" represents A or G, "M" represents A or C; "W" represents A or T; "Y" represents C or T; "S" represents C or G; "K" represents G or T; "V" represents A, C or G; "H" represents A, C, or T; "D" represents A, G, or T; "B" represents C, G, or T; and "N" represents A, G, C, or T.

[0012] Figure 2 shows a polypeptide sequence encoded by the nucleic acid of SEQ. ID. NO: 1.

[0013] Figures 3A and 3C depict tissue expression profiles for PLA2G1B and Figures 3B and 3D show expanded profiles of Figures 3A and 3C, respectively.

[0014] Figures 4A to 4L show differential gene expression of PLA2G1B in metabolically-linked tissues, such as liver, fat pads, skeletal muscle, hypothalamus, pancreas, and stomach tissues from were analyzed following normal feeding or overnight fasting conditions. Studies were typically performed on group A (healthy), B (insulin resistant) and C animals (Diabetic/Obese), as group D animals (Diabetic/Obese) developed decompensated diabetes when their pancreas failed, leading to rapid death. In addition, the Figures contain data relating to blood glucose, plasma insulin, body weight, and body fat from the animals as compared to gene expression using t-test analysis. Figure 4A shows PLA2G1B expression in the hypothalamus in group C fasted animals as compared to group A fasted animals and group B fasted animals. Figure 4B shows hypothalamus PLA2G1B expression in group A animals that were fed normally versus fasted group A animals. Figure 4C shows hypothalamus PLA2G1B expression

in fasted animals versus body weight. Figure 4D shows hypothalamus PLA2G1B expression in fasted animals versus plasma insulin levels. Figure 4E shows expression in A fasted animals as compared to C fasted and B fasted animals. Figure 4F shows expression in A fed group versus C fed group. Figures 4G, 4H and 4I show gene expression in fasted animals versus body weight, insulin and glucose. Figure 4J shows liver PLA2G1B expression in fed animals versus body weight (p=0.013). Figure 4K shows pancreatic PLA2G1B expression in control versus energy-restricted groups. Figure 4L shows PLA2G1B expression in the fasted animals versus the fed animals.

[0015] Figure 5A shows a nucleotide sequence alignment for human PLA2G1B and related sequences from mouse, rat, and *P. obesus* (sand rat). Figure 5B shows an amino acid sequence alignment between human PLA2G1B and related sequences from mouse, rat, and *P. obesus*. The human PLA2G1B amino acid sequence in Figure 5B has 148 amino acids and the mouse, rat, and *P. obesus* sequences have 146 amino acids. The human PLA2G1B amino acid sequence is 78% identical to the mouse sequence, 76% identical to the rat sequence, and 76% identical to the *P. obesus* sequence. The mouse sequence is 88% identical to the rat sequence and 77% identical to the *P. obesus* sequence, and the rat sequence is 80% identical to the *P. obesus* sequence.

Detailed Description

[0016] It has been discovered that polymorphic variants in or near a gene on chromosome 12 encoding a phospholipase are associated with fat deposition in the abdomen and trunk region of subjects. Individuals having increased fat deposition in this area are at risk of developing metabolic conditions (*e.g.*, diabetes and obesity) and cardiovascular conditions (*e.g.*, hypertension). Thus, methods for detecting genetic determinants for fat deposition can lead to early diagnosis of a predisposition to these conditions (*e.g.*, hyperinsulinaemia, hypertension, glucose intolerance (that is, IGT or diabetes), dyslipidemia, hypercoagulability and microalbuminuria) and early prescription of preventative measures. Thus, methods for detecting genetic determinants for fat deposition can lead to early diagnosis of a predisposition to these conditions and early prescription of preventative measures.

Central Fat Deposition and Associated Conditions

[0017] Many individuals considered as having increased central fat deposition are also considered obese according to BMI, weight-for-height charts, or body fat measurements. Obesity is generally understood as a condition where fat content in an individual is above a predetermined level. For example, the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) of the National Institutes of Health (NIH, *see* http address www.nih.gov/health/nutrit/pubs/ statobes.htm) define individuals having a body mass index (BMI) of 25 to 29.9 kg/m^2 as being overweight and individuals having a BMI of 30 kg/m^2 or greater as being obese.

[0018] Increased central fat levels also have been linked to the metabolic syndrome, which includes the coexistence or one or more life threatening medical conditions such as metabolic conditions (*e.g.*, diabetes and obesity) and cardiovascular conditions (*e.g.*, myocardial infarction and hypertension). For example, cardiovascular mortality was assessed in 3,606 subjects from the Botnia study (a large-scale study of type 2 diabetes begun in Finland in 1990) with a median follow-up of 6.9 years. In women and men, respectively, the metabolic syndrome was recorded in 10 and 15% of subjects with normal glucose tolerance, 42 and 64% of those with IFG/IGT, and 78 and 84% of those with type 2 diabetes. The risk for coronary heart disease and stroke was increased threefold in subjects with the syndrome, and cardiovascular mortality was markedly increased (12.0% in subjects with the syndrome versus 2.2% in those without; $P < 0.001$) (Zimmet, *et al.* (2001) *Nature* 414: 782-787). Thus, determining a predisposition to fat deposition, and specifically central fat deposition, is useful for determining whether a person should be considered for being placed on a preventative regimen for reducing fat, thereby reducing the probability that the person develops one or more conditions linked to fat deposition.

[0019] The term "fat deposition" as used herein refers to fat content in an individual as well as processes in which fat is deposited in certain locations of an individual. The term "central fat deposition" as used herein refers to fat around the trunk and waist of an individual that is above a predetermined level or average in a population. The central region may be defined as the region extending from the superior surface of the second lumbar vertebra extending inferiorly to the inferior surface of the fourth lumbar vertebra and laterally to the inner aspect of the ribcage. Fat deposition can be measured as a quantity at one time point or a quantity over a series of time points, for example, and fat deposition can be quantified or estimated using a number of procedures described hereafter. Fat is composed of adipose cells deposited below the skin (*i.e.*, subcutaneous adipose cells) and/or deeper within an individual's body (*i.e.*, visceral adipose cells). Adipose cells are often connective tissue cells specialized for synthesis and storage of fat. Such cells often contain globules of triglycerides where the nucleus is generally displaced to one side of the globule and the cytoplasm is visualized as a thin line around the fat droplet. Provided herein are methods for detecting predisposition to overall adipose cell deposition in a subject (*i.e.*, includes subcutaneous adipose cells and visceral adipose cells), as well as methods for distinguishing between a predisposition to subcutaneous adipose cell deposition and a predisposition to visceral adipose cell deposition.

[0020] Fat deposition may be quantified in a number of manners (*see, e.g.*, Wajchenberg, *Endocrine Rev. 21(6)*: 697-738 (2000)). For example, caliper measurements of skinfold thickness in defined areas of the body have been utilized to differ between different kinds of regional fat (Nordhamn, *et al.*, *Int. J. Obes. Relat. Metab. Disord. 24(5):* 652-7 (2000)). Waist and hip measurements using tape measures are commonly utilized indices of central fat (Lundgren *et al.*, *Int. J. Obes., 13(4):* 413-23 (1989); Ohlson *et al.*, *Diabetes 34(10):* 1055-8 (1985)), and sagittal abdominal diameter is measured by some researchers

for quantifying central fat. Also, computed tomography and X-ray based methods have been utilized to quantify central fat content. Dual x-ray absorbtiometry (DEXA) is relatively fast and inexpensive and yields reliable estimations of body composition (fat mass/lean mass /bone) with reproducibility. DEXA measurements and waist and hip measurements were utilized for quantifying central fat in Example 1. Magnetic resonance imaging (MRI) and computed tomography procedures can be used to distinguish between visceral fat deposition and subcutaneous fat deposition (*see e.g.* Wajchenberg, *supra*).

[0021] Thus, fat deposition can be expressed in terms of any units used for quantifying fat content. Fat deposition can be expressed in terms of total fat content in an individual or region of an individual (grams or percentage of total weight of an individual), visceral fat content in an individual or region of an individual (grams, percentage of total weight of an individual, or percentage of total fat in an individual), and subcutaneous fat content in an individual or region of an individual (grams, percentage of total weight of an individual, or percentage of total fat in an individual). Each of these expressions of fat deposition can be measured or quantified at a single point in time or over two or more points in time.

[0022] Fat deposition also can be expressed in terms of "increased fat deposition" (also referred to as "higher fat deposition" and "at increased risk for fat deposition"), which is relative to average fat deposition in a population. In a distribution of fat deposition across a population (expressed in any of the units of measure described herein), individuals having increased fat deposition are sometimes represented in the upper 40% or upper 30% of the population, often in the upper 25%, upper 20%, upper 15%, and upper 10% of the population, and sometimes in the upper 5% of the population. Also, individuals having increased fat deposition can be characterized as having waist/hip ratios of 1.01 or more for males and 0.91 or more for females. In addition, men or women having a BMI between 25 and 30 or between about 1335 and about 2050 grams of central fat are typically considered overweight, and individuals having a BMI over 30 or over about 2050 grams of central fat are normally considered obese (*e.g.*, grams of central fat can be determined by DEXA, as described above). Also, "leanness" or "decreased fat deposition" (also referred to as "lower fat deposition" and "at decreased risk for fat deposition") are terms that refer to fat deposition and are also relative to average fat deposition in a population. In a distribution of fat deposition across a population, lean individuals are sometimes represented in the lower 40% or lower 30% of the population, often in the lower 25%, lower 20%, lower 15%, and lower 10% of the population, and sometimes in the lower 5% of the population. Also, lean individuals can be characterized as having waist/hip ratios of 1.00 or less for males and 0.90 or less for females. In addition, men or women having a BMI of 24 or less or less than about 1334 grams of central fat are normally considered lean.

[0023] The term "metabolic condition" as used herein refers to a disease, disorder, or state involving increased or decreased metabolites relative to a population average. Examples of metabolic disorders include but are not limited to diabetes, obesity, anorexia nervosa, cachexia, and lipid disorders.

[0024] The term "NIDDM" as used herein refers to non-insulin-dependent diabetes mellitus or Type 2 diabetes (the two terms are used interchangeably throughout this document). NIDDM refers to an insulin-related disorder in which there is a relative disparity between endogenous insulin production and insulin requirements, leading to elevated hepatic glucose production, elevated blood glucose levels, inappropriate insulin secretion, and peripheral insulin resistance.

[0025] The term "cardiovascular condition" as used herein refers to a disease, disorder, or state involving the cardiovascular system, *e.g.*, the heart, the blood vessels, and/or the blood. A cardiovascular disorder can be caused by an imbalance in arterial pressure, a malfunction of the heart, or an occlusion of a blood vessel (*e.g.*, by a thrombus). Other examples of cardiovascular disorders include but are not limited to hypertension, atherosclerosis, coronary artery spasm, coronary artery disease, arrhythmias, heart failure, including but not limited to, cardiac hypertrophy, left-sided heart failure, and right-sided heart failure; ischemic heart disease, including but not limited to angina pectoris, myocardial infarction, chronic ischemic heart disease, and sudden cardiac death; hypertensive heart disease, including but not limited to, systemic (left-sided) hypertensive heart disease and pulmonary (right-sided) hypertensive heart disease; valvular heart disease, including but not limited to, valvular degeneration caused by calcification, such as calcification of a congenitally bicuspid aortic valve, and mitral annular calcification, and myxomatous degeneration of the mitral valve (mitral valve prolapse), rheumatic fever and rheumatic heart disease, infective endocarditis, and noninfected vegetations, such as nonbacterial thrombotic endocarditis and endocarditis of systemic lupus erythematosus (Libman-Sacks disease), carcinoid heart disease, and complications of artificial valves; myocardial disease, including but not limited to dilated cardiomyopathy, hypertrophic cardiomyopathy, restrictive cardiomyopathy, and myocarditis; pericardial disease, including but not limited to, pericardial effusion and hemopericardium and pericarditis, including acute pericarditis and healed pericarditis, and rheumatoid heart disease; neoplastic heart disease, including but not limited to, primary cardiac tumors, such as myxoma, lipoma, papillary fibroblastoma, rhabdomyoma, and sarcoma, and cardiac effects of noncardiac neoplasms; congenital heart disease, including but not limited to, left-to-right shunts (late cyanosis, such as atrial septal defect, ventricular septal defect, patent ductus arteriosis, and atrioventricular septal defect, right-to-left shunts), early cyanosis (*e.g.*, tetralogy of fallot, transposition of great arteries, truncus arteriosis, tricuspid atresia, and total anomalous pulmonary venous connection), obstructive congenital anomalies (*e.g.*, coarctation of aorta, pulmonary stenosis and atresia, and aortic stenosis and atresia), disorders involving cardiac transplantation, and congestive heart failure.

Polymorphic Variants Associated with Fat Deposition and Related Conditions

[0026] A genetic analysis provided herein linked fat deposition with polymorphic variants of a nucleotide sequence located on chromosome twelve that encodes a phospholipase A2 polypeptide

designated PLA2G1B. An additional genetic analysis provided herein linked NIDDM with a polymorphic variant of a nucleotide sequence located on chromosome twelve that encodes a phospholipase A2, group IB designated PLA2G1B. As used herein, the term "polymorphic site" refers to a region in a nucleic acid at which two or more alternative nucleotide sequences are observed in a significant number of nucleic acid samples from a population of individuals. A polymorphic site may be a nucleotide sequence of two or more nucleotides, an inserted nucleotide or nucleotide sequence, a deleted nucleotide or nucleotide sequence, or a microsatellite, for example. A polymorphic site that is two or more nucleotides in length may be 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or more, 20 or more, 30 or more, 50 or more, 75 or more, 100 or more, 500 or more, or about 1000 nucleotides in length, where all or some of the nucleotide sequences differ within the region. A polymorphic site is often one nucleotide in length, which is referred to herein as a "single nucleotide polymorphism" or a "SNP."

[0027] Where there are two, three, or four alternative nucleotide sequences at a polymorphic site, each nucleotide sequence is referred to as a "polymorphic variant." Where two polymorphic variants exist, for example, the polymorphic variant represented in a minority of samples from a population is sometimes referred to as a "minor allele" and the polymorphic variant that is more prevalently represented is sometimes referred to as a "major allele." Many organisms possess a copy of each chromosome (*e.g.*, humans), and those individuals who possess two major alleles or two minor alleles are often referred to as being "homozygous" with respect to the polymorphism and those individuals who possess one major allele and one minor allele are normally referred to as being "heterozygous" with respect to the polymorphism. Individuals who are homozygous with respect to one allele are sometimes predisposed to a different phenotype as compared to individuals who are heterozygous or homozygous with respect to another allele. As used herein, the term "phenotype" refers to a trait which can be compared between individuals, such as presence or absence of a condition, a visually observable difference in appearance between individuals, metabolic variations, physiological variations, variations in the function of biological molecules, and the like. Examples of phenotypes are fat deposition, obesity, and diabetes.

[0028] Researchers sometimes report a polymorphic variant in a database without determining whether the variant is represented in a significant fraction of a population. Because a subset of these reported polymorphismic variants are not represented in a statistically significant portion of the population, some of them are sequencing errors and/or not biologically relevant. Thus, it is often not known whether a reported polymorphic variant is statistically significant or biologically relevant until the presence of the variant is detected in a population of individuals and the frequency of the variant is determined. Methods for detecting a polymorphic variant in a population are described herein, specifically in Example 2. A polymorphic variant is statistically significant and often biologically relevant if it is represented in 5% or more of a population, sometimes 10% or more, 15% or more, or 20%

or more of a population, and often 25% or more, 30% or more, 35% or more, 40% or more, 45% or more, or 50% or more of a population.

[0029] A polymorphic variant may be detected on either or both strands of a double-stranded nucleic acid. Also, a polymorphic variant may be located within an intron or exon of a gene or within a portion of a regulatory region such as a promoter, a 5′ untranslated region (UTR), a 3′ UTR, and in DNA (*e.g.*, genomic DNA (gDNA) and complementary DNA (cDNA)), RNA (*e.g.*, mRNA, tRNA, and rRNA), or a polypeptide. Polymorphic variations may or may not result in detectable differences in gene expression, polypeptide structure, or polypeptide function.

[0030] In the genetic analysis that associated polymorphic variations in PLA2G1B with fat deposition, samples from individuals in a population of twin pairs were genotyped, although other populations could be subjected to analysis. The term "genotyped" as used herein refers to a process for determining a genotype of one or more individuals, where a "genotype" is a representation of polymorphic variants in a population. Fat deposition was quantified in the central region of individuals in the study group, and SNPs were identified at positions 7328 and 9182 in the PLA2G1B nucleotide sequence represented by SEQ ID NO:1. It was determined that 84% of the individuals tested in the genetic analysis had a guanine at position 7328 and 16% of the individuals had an adenine at this position. At position 9182, 85% of the individuals had a thymine and 15% of the individuals had a guanine. It was determined that a guanine at position 7328 or a thymine at position 9182 were individually associated with central fat deposition, and the presence of an adenine at position 7328 or a guanine at position 9182 were individually associated with leanness.

[0031] In the genetic analysis that associated polymorphic variations in PLA2G1B with NIDDM, samples from individuals in a population of with NIDDM and without NIDDM were genotyped. A SNP was identified at position 7256 in the PLA2G1B nucleotide sequence represented by SEQ ID NO:1. It was determined that 93% of female controls tested in the genetic analysis had a thymine at position 7256 and 7% of the individuals had a cytosine at this position, while 92% of female cases tested in the genetic analysis had a thymine at position 7256 and 8% of the individuals had a cytosine at this position. It was also determined that 95% of male controls tested in the genetic analysis had a thymine at position 7256 and 5% of the individuals had a cytosine at this position, while 90% of male cases tested in the genetic analysis had a thymine at position 7256 and 10% of the individuals had a cytosine at this position. It was determined that a cytosine at position 7256 was individually associated with NIDDM, and the presence of a thymine at position 7256 was individually associated with not having NIDDM.

[0032] Furthermore, a genotype or polymorphic variant may be expressed in terms of a "haplotype," which as used herein refers to two or more polymorphic variants occurring within genomic DNA in a group of individuals within a population. For example, two SNPs may exist within a gene where each SNP position includes a cytosine variation and an adenine variation. Certain individuals in a population

may carry one allele (heterozygous) or two alleles (homozygous) having the gene with a cytosine at each SNP position. As the two cytosines corresponding to each SNP in the gene travel together on one or both alleles in these individuals, the individuals can be characterized as having a cytosine/cytosine haplotype with respect to the two SNPs in the gene.

[0033] Also, the genetic analysis identified haplotypes associated with lower risk of fat deposition. In particular, presence of a haplotype represented by TTAG or GTAG at positions 4050, 7256, 7328, and 9182, respectively, in the PLA2G1B sequence represented by SEQ ID NO:1 were associated with leanness. As used herein, a "haplotype" refers to a combination of polymorphic variations in a defined region within a genetic locus on one of the chromosomes in a chromosome pair.

Additional Polymorphic Variants Associated with Fat Deposition and Related Disorders

[0034] Also provided is a method for identifying polymorphic variants proximal to an incident, founder polymorphic variant associated with fat deposition, obesity and NIDDM. Thus, featured herein are methods for identifying a polymorphic variation associated with fat deposition or NIDDM that is proximal to an incident polymorphic variation associated with fat deposition or NIDDM, which comprises identifying a polymorphic variant proximal to the incident polymorphic variant associated with fat deposition of NIDDM, where the incident polymorphic variant is in a PLA2G1B nucleotide sequence. The PLA2G1B nucleotide sequence often comprises a polynucleotide sequence selected from the group consisting of (a) a polynucleotide sequence set forth in SEQ ID NO: 1; (b) a polynucleotide sequence that encodes a polypeptide having an amino acid sequence encoded by a nucleotide sequence set forth as SEQ ID NO: 1; or (c) a polynucleotide sequence that encodes a polypeptide having an amino acid sequence that is 90% identical to an amino acid sequence encoded by a nucleotide sequence set forth in SEQ ID NO: 1 or a polynucleotide sequence 90% identical to the polynucleotide sequence of SEQ ID NO:1. The presence or absence of an association of the proximal polymorphic variant with fat deposition or NIDDM then is determined using a known association method, such as a method described in the Examples hereafter. In an embodiment, the incident polymorphic variant is at position 7256, 7328, or 9182 of SEQ ID NO: 1. In another embodiment, the proximal polymorphic variant identified sometimes is a publicly disclosed polymorphic variant, which for example, sometimes is published in a publicly available database. In other embodiments, the polymorphic variant identified is not publicly disclosed and is discovered using a known method, including, but not limited to, sequencing a region surrounding the incident polymorphic variant in a group of nucleic samples. Thus, multiple polymorphic variants proximal to an incident polymorphic variant are associated with fat deposition and NIDDM using this method.

[0035] The proximal polymorphic variant often is identified in a region surrounding the incident polymorphic variant. In certain embodiments, this surrounding region is about 50 kb flanking the first

polymorphic variant (*e.g.* about 50 kb 5' of the first polymorphic variant and about 50 kb 3' of the first polymorphic variant), and the region sometimes is composed of shorter flanking sequences, such as flanking sequences of about 40 kb, about 30 kb, about 25 kb, about 20 kb, about 15 kb, about 10 kb, about 7 kb, about 5 kb, or about 2 kb 5' and 3' of the incident polymorphic variant. In other embodiments, the region is composed of longer flanking sequences, such as flanking sequences of about 55 kb, about 60 kb, about 65 kb, about 70 kb, about 75 kb, about 80 kb, about 85 kb, about 90 kb, about 95 kb, or about 100 kb 5' and 3' of the incident polymorphic variant.

[0036] In certain embodiments, polymorphic variants associated with fat deposition or NIDDM are identified iteratively. For example, a first proximal polymorphic variant is associated with fat deposition using the methods described above and then another polymorphic variant proximal to the first proximal polymorphic variant is identified (*e.g.*, publicly disclosed or discovered) and the presence or absence of an association of one or more other polymorphic variants proximal to the first proximal polymorphic variant with fat deposition or NIDDM is determined.

[0037] The methods described herein are useful for identifying or discovering additional polymorphic variants that may be used to further characterize a gene, region or loci associated with a condition, a disease (*e.g.*, fat deposition or NIDDM), or a disorder. For example, allelotyping or genotyping data from the additional polymorphic variants may be used to identify a functional mutation or a region of linkage disequilibrium.

[0038] In certain embodiments, polymorphic variants identified or discovered within a region comprising the first polymorphic variant associated with fat deposition or NIDDM are genotyped using the genetic methods and sample selection techniques described herein, and it can be determined whether those polymorphic variants are in linkage disequilibrium with the first polymorphic variant. The size of the region in linkage disequilibrium with the first polymorphic variant also can be assessed using these genotyping methods. Thus, provided herein are methods for determining whether a polymorphic variant is in linkage disequilibrium with a first polymorphic variant associated with fat deposition or NIDDM, and such information can be used in prognosis methods described herein.

Isolated PLA2G1B Nucleic Acids and Variants Thereof

[0039] Featured herein are isolated PLA2G1B nucleic acids, which include the nucleic acid having the nucleotide sequence of SEQ ID NO:1, PLA2G1B nucleic acid variants, and substantially identical nucleic acids to the foregoing. Nucleotide sequences of the PLA2G1B nucleic acids are sometimes referred to herein as "PLA2G1B nucleotide sequences." A "PLA2G1B nucleic acid variant" refers to one allele that may have different polymorphic variations as compared to another allele in another subject or

the same subject. A polymorphic variation in the PLA2G1B nucleic acid variant may be represented on one or both strands in a double-stranded nucleic acid or on one chromosomal complement (heterozygous) or both chromosomal complements (homozygous)). A PLA2G1B nucleic acid may comprise one or more of the following polymorphic variations: a thymine or a cytosine at position 7256 of SEQ ID NO:1 in a strand, or an adenine or guanine in a complementary strand;an adenine or guanine at position 7328 of SEQ ID NO:1 in a strand, or a thymine or cytosine in a complementary strand; or a guanine or thymine at position 9182 of SEQ ID NO:1 in a strand, or a cytosine or adenine in a complementary strand; presence of GTGT, TTGT, TTAG, GCGT, or GTAG at positions 4050, 7256, 7328, and 9182 of SEQ ID NO:1, respectively, in a strand, or presence of CACA, AACA, AATC, CGCA, or CATC in a complementary strand.

[0040] As used herein, the term "nucleic acid" includes DNA molecules (*e.g.*, a complementary DNA (cDNA) and genomic DNA (gDNA)) and RNA molecules (*e.g.*, mRNA, rRNA, siRNA and tRNA) and analogs of DNA or RNA, for example, by use of nucleotide analogs. The nucleic acid molecule can be single-stranded and it is often double-stranded. The term "isolated or purified nucleic acid" refers to nucleic acids that are separated from other nucleic acids present in the natural source of the nucleic acid. For example, with regard to genomic DNA, the term "isolated" includes nucleic acids which are separated from the chromosome with which the genomic DNA is naturally associated. An "isolated" nucleic acid is often free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and/or 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated nucleic acid molecule can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of 5' and/or 3' nucleotide sequences which flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. As used herein, the term "PLA2G1B gene" refers to a nucleotide sequence that encodes a PLA2G1B polypeptide.

[0041] Also included herein are nucleic acid fragments. These fragments are typically a nucleotide sequence identical to a nucleotide sequence in SEQ ID NO:1, a nucleotide sequence substantially identical to a nucleotide sequence in SEQ ID NO:1, or a nucleotide sequence that is complementary to the foregoing. The nucleic acid fragment may be identical, substantially identical or homologous to a nucleotide sequence in an exon or an intron in SEQ ID NO:1 and may encode a domain or part of a domain of a PLA2G1B polypeptide. Sometimes, the fragment will comprises one or more of the polymorphic variations described herein as being associated with increased fat deposition or increased risk of developing NIDDM. The nucleic acid fragment is often 50, 100, or 200 or fewer base pairs in length, and is sometimes about 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500,

2000, 3000, 4000, 5000, 10000, or 12000 base pairs in length. A nucleic acid fragment that is complementary to a nucleotide sequence identical or substantially identical to the nucleotide sequence of SEQ ID NO:1 and hybridizes to such a nucleotide sequence under stringent conditions is often referred to as a "probe." Nucleic acid fragments often include one or more polymorphic sites, or sometimes have an end that is adjacent to a polymorphic site as described hereafter.

[0042] An example of a nucleic acid fragment is an oligonucleotide. As used herein, the term "oligonucleotide" refers to a nucleic acid comprising about 8 to about 50 covalently linked nucleotides, often comprising from about 8 to about 35 nucleotides, and more often from about 10 to about 25 nucleotides. The backbone and nucleotides within an oligonucleotide may be the same as those of naturally occurring nucleic acids, or analogs or derivatives of naturally occurring nucleic acids, provided that oligonucleotides having such analogs or derivatives retain the ability to hybridize specifically to a nucleic acid comprising a targeted polymorphism. Oligonucleotides described herein may be used as hybridization probes or as components of diagnostic assays, for example, as described herein.

[0043] Oligonucleotides are typically synthesized using standard methods and equipment, such as the ABI™3900 High Throughput DNA Synthesizer and the EXPEDITE™ 8909 Nucleic Acid Synthesizer, both of which are available from Applied Biosystems (Foster City, CA). Analogs and derivatives are exemplified in U.S. Pat. Nos. 4,469,863; 5,536,821; 5,541,306; 5,637,683; 5,637,684; 5,700,922; 5,717,083; 5,719,262; 5,739,308; 5,773,601; 5,886,165; 5,929,226; 5,977,296; 6,140,482; WO 00/56746; WO 01/14398, and related publications. Methods for synthesizing oligonucleotides comprising such analogs or derivatives are disclosed, for example, in the patent publications cited above and in U.S. Pat. Nos. 5,614,622; 5,739,314; 5,955,599; 5,962,674; 6,117,992; in WO 00/75372; and in related publications.

[0044] Oligonucleotides may also be linked to a second moiety. The second moiety may be an additional nucleotide sequence such as a tail sequence (*e.g.*, a polyadenosine tail), an adaptor sequence (*e.g.*, phage M13 universal tail sequence), and others. Alternatively, the second moiety may be a non-nucleotide moiety such as a moiety which facilitates linkage to a solid support or a label to facilitate detection of the oligonucleotide. Such labels include, without limitation, a radioactive label, a fluorescent label, a chemiluminescent label, a paramagnetic label, and the like. The second moiety may be attached to any position of the oligonucleotide, provided the oligonucleotide can hybridize to the nucleic acid comprising the polymorphism.

Uses for Nucleic Acid Sequence

[0045] Nucleic acid coding sequences depicted in SEQ ID NO: 1 may be used for diagnostic purposes for detection and control of polypeptide expression. Also, included herein are oligonucleotide sequences such as antisense RNA, small-interfering RNA (siRNA) and DNA molecules and ribozymes that function

to inhibit translation of a polypeptide. Antisense techniques and RNA interference techniques are known in the art and are described herein.

[0046] Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. The mechanism of ribozyme action involves sequence specific hybridization of the ribozyme molecule to complementary target RNA, followed by a endonucleolytic cleavage. Ribozymes may be engineered hammerhead motif ribozyme molecules that specifically and efficiently catalyze endonucleolytic cleavage of RNA sequences corresponding to or complementary to the nucleotide sequences set forth in SEQ ID NO: 1. Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences, GUA, GUU and GUC. Once identified, short RNA sequences of between fifteen (15) and twenty (20) ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for predicted structural features such as secondary structure that may render the oligonucleotide sequence unsuitable. The suitability of candidate targets may also be evaluated by testing their accessibility to hybridization with complementary oligonucleotides, using ribonuclease protection assays.

[0047] Antisense RNA and DNA molecules, siRNA and ribozymes may be prepared by any method known in the art for the synthesis of RNA molecules. These include techniques for chemically synthesizing oligodeoxyribonucleotides well known in the art such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding the antisense RNA molecule. Such DNA sequences may be incorporated into a wide variety of vectors which incorporate suitable RNA polymerase promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense cDNA constructs that synthesize antisense RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines.

[0048] DNA encoding a polypeptide also may have a number of uses for the diagnosis of diseases, including fat deposition or NIDDM, resulting from aberrant expression of PLA2G1B. For example, the nucleic acid sequence may be used in hybridization assays of biopsies or autopsies to diagnose abnormalities of expression or function (*e.g.*, Southern or Northern blot analysis, in situ hybridization assays).

[0049] In addition, the expression of a polypeptide during embryonic development may also be determined using nucleic acid encoding the polypeptide. As addressed, *infra*, production of functionally impaired polypeptide is the cause of various disease states, including fat deposition or NIDDM. *In situ* hybridizations using polypeptide as a probe may be employed to predict problems related to obesity or NIDDM. Further, as indicated, *infra*, administration of human active polypeptide, recombinantly produced as described herein, may be used to treat disease states related to functionally impaired polypeptide. Alternatively, gene therapy approaches may be employed to remedy deficiencies of functional polypeptide or to replace or compete with dysfunctional polypeptide.

Expression Vectors, Host Cells, and Genetically Engineered Cells

[0050] Provided herein are nucleic acid vectors, often expression vectors, which contain a PLA2G1B nucleic acid. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked and can include a plasmid, cosmid, or viral vector. The vector can be capable of autonomous replication or it can integrate into a host DNA. Viral vectors may include replication defective retroviruses, adenoviruses and adeno-associated viruses for example.

[0051] A vector can include a PLA2G1B nucleic acid in a form suitable for expression of the nucleic acid in a host cell. The recombinant expression vector typically includes one or more regulatory sequences operatively linked to the nucleic acid sequence to be expressed. The term "regulatory sequence" includes promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Regulatory sequences include those that direct constitutive expression of a nucleotide sequence, as well as tissue-specific regulatory and/or inducible sequences. The design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of polypeptide desired, and the like. Expression vectors can be introduced into host cells to produce PLA2G1B polypeptides, including fusion polypeptides, encoded by PLA2G1B nucleic acids.

[0052] Recombinant expression vectors can be designed for expression of PLA2G1B polypeptides in prokaryotic or eukaryotic cells. For example, PLA2G1B polypeptides can be expressed in *E. coli*, insect cells (*e.g.*, using baculovirus expression vectors), yeast cells, or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology 185*, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

[0053] Expression of polypeptides in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion polypeptides. Fusion vectors add a number of amino acids to a polypeptide encoded therein, usually to the amino terminus of the recombinant polypeptide. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant polypeptide; 2) to increase the solubility of the recombinant polypeptide; and 3) to aid in the purification of the recombinant polypeptide by acting as a ligand in affinity purification. Often, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant polypeptide to enable separation of the recombinant polypeptide from the fusion moiety subsequent to purification of the fusion polypeptide. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith, D. B. and Johnson, K. S., *Gene 67:* 31-40 (1988)), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-

transferase (GST), maltose E binding polypeptide, or polypeptide A, respectively, to the target recombinant polypeptide.

[0054] Purified fusion polypeptides can be used in screening assays and to generate antibodies specific for PLA2G1B polypeptides. In a therapeutic embodiment, fusion polypeptide expressed in a retroviral expression vector is used to infect bone marrow cells that are subsequently transplanted into irradiated recipients. The pathology of the subject recipient is then examined after sufficient time has passed (*e.g.*, six (6) weeks).

[0055] Expressing the polypeptide in host bacteria with an impaired capacity to proteolytically cleave the recombinant polypeptide is often used to maximize recombinant polypeptide expression (Gottesman, S., *Gene Expression Technology: Methods in Enzymology, Academic Press, San Diego, California 185:* 119-128 (1990)). Another strategy is to alter the nucleotide sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al., Nucleic Acids Res. 20:* 2111-2118 (1992)). Such alteration of nucleotide sequences can be carried out by standard DNA synthesis techniques.

[0056] When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. Recombinant mammalian expression vectors are often capable of directing expression of the nucleic acid in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Non-limiting examples of suitable tissue-specific promoters include an albumin promoter (liver-specific; Pinkert *et al., Genes Dev. 1:* 268-277 (1987)), lymphoid-specific promoters (Calame and Eaton, *Adv. Immunol. 43:* 235-275 (1988)), promoters of T cell receptors (Winoto and Baltimore, *EMBO J. 8:* 729-733 (1989)) promoters of immunoglobulins (Banerji *et al., Cell 33:* 729-740 (1983); Queen and Baltimore, *Cell 33:* 741-748 (1983)), neuron-specific promoters (*e.g.*, the neurofilament promoter; Byrne and Ruddle, *Proc. Natl. Acad. Sci. USA 86:* 5473-5477 (1989)), pancreas-specific promoters (Edlund *et al., Science 230:* 912-916 (1985)), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are sometimes utilized, for example, the murine hox promoters (Kessel and Gruss, *Science 249:* 374-379 (1990)) and the α-fetopolypeptide promoter (Campes and Tilghman, *Genes Dev. 3:* 537-546 (1989)).

[0057] A PLA2G1B nucleic acid may also be cloned into an expression vector in an antisense orientation. Regulatory sequences (*e.g.*, viral promoters and/or enhancers) operatively linked to a PLA2G1B nucleic acid cloned in the antisense orientation can be chosen for directing constitutive, tissue specific or cell type specific expression of antisense RNA in a variety of cell types. Antisense expression vectors can be in the form of a recombinant plasmid, phagemid or attenuated virus. For a discussion of

the regulation of gene expression using antisense genes see Weintraub, H. et al., Antisense RNA as a molecular tool for genetic analysis, *Reviews - Trends in Genetics*, Vol. 1(1) (1986).

[0058] Also provided herein are host cells that include a PLA2G1B nucleic acid within a recombinant expression vector or PLA2G1B nucleic acid sequence fragments which allow it to homologously recombine into a specific site of the host cell genome. The terms "host cell" and "recombinant host cell" are used interchangeably herein. Such terms refer not only to the particular subject cell but rather also to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein. A host cell can be any prokaryotic or eukaryotic cell. For example, a PLA2G1B polypeptide can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

[0059] Vectors can be introduced into host cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation.

[0060] A host cell provided herein can be used to produce (*i.e.*, express) a PLA2G1B polypeptide. Accordingly, further provided are methods for producing a PLA2G1B polypeptide using the host cells of the invention. In one embodiment, the method includes culturing host cells into which a recombinant expression vector encoding a PLA2G1B polypeptide has been introduced in a suitable medium such that a PLA2G1B polypeptide is produced. In another embodiment, the method further includes isolating a PLA2G1B polypeptide from the medium or the host cell.

[0061] Also provided are cells or purified preparations of cells which include a PLA2G1B transgene, or which otherwise misexpress PLA2G1B polypeptide. Cell preparations can consist of human or non-human cells, *e.g.*, rodent cells, *e.g.*, mouse or rat cells, rabbit cells, or pig cells. In preferred embodiments, the cell or cells include a PLA2G1B transgene (*e.g.*, a heterologous form of a PLA2G1B such as a human gene expressed in non-human cells). The PLA2G1B transgene can be misexpressed, *e.g.*, overexpressed or underexpressed. In other preferred embodiments, the cell or cells include a gene which misexpress an endogenous PLA2G1B polypeptide (*e.g.*, expression of a gene is disrupted, also known as a knockout). Such cells can serve as a model for studying disorders which are related to mutated or mis-expressed PLA2G1B alleles or for use in drug screening. Also provided are human cells (*e.g.*, a hematopoietic stem cells) transformed with a PLA2G1B nucleic acid.

[0062] Also provided are cells or a purified preparation thereof (*e.g.*, human cells) in which an endogenous PLA2G1B nucleic acid is under the control of a regulatory sequence that does not normally control the expression of the endogenous PLA2G1B gene. The expression characteristics of an endogenous gene within a cell (*e.g.*, a cell line or microorganism) can be modified by inserting a heterologous DNA regulatory element into the genome of the cell such that the inserted regulatory element is operably linked to the endogenous PLA2G1B gene. For example, an endogenous PLA2G1B gene (*e.g.*, a gene which is "transcriptionally silent," not normally expressed, or expressed only at very low levels) may be activated by inserting a regulatory element which is capable of promoting the expression of a normally expressed gene product in that cell. Techniques such as targeted homologous recombinations, can be used to insert the heterologous DNA as described in, *e.g.*, Chappel, US 5,272,071; WO 91/06667, published on May 16, 1991.

Transgenic Animals

[0063] Non-human transgenic animals that express a heterologous PLA2G1B polypeptide (*e.g.*, expressed from a PLA2G1B nucleic acid isolated from another organism) can be generated. Such animals are useful for studying the function and/or activity of a PLA2G1B polypeptide and for identifying and/or evaluating modulators of PLA2G1B nucleic acid and PLA2G1B polypeptide activity. As used herein, a "transgenic animal" is a non-human animal such as a mammal (*e.g.*, a non-human primate such as chimpanzee, baboon, or macaque; an ungulate such as an equine, bovine, or caprine; or a rodent such as a rat, a mouse, or an Israeli sand rat), a bird (*e.g.*, a chicken or a turkey), an amphibian (*e.g.*, a frog, salamander, or newt), or an insect (*e.g.*, *drosophila melanogaster*), in which one or more of the cells of the animal includes a PLA2G1B transgene. A transgene is exogenous DNA or a rearrangement (*e.g.*, a deletion of endogenous chromosomal DNA) that is often integrated into or occurs in the genome of cells in a transgenic animal. A transgene can direct expression of an encoded gene product in one or more cell types or tissues of the transgenic animal, and other transgenes can reduce expression (*e.g.*, a knockout). Thus, a transgenic animal can be one in which an endogenous PLA2G1B gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal (*e.g.*, an embryonic cell of the animal) prior to development of the animal.

[0064] Intronic sequences and polyadenylation signals can also be included in the transgene to increase expression efficiency of the transgene. One or more tissue-specific regulatory sequences can be operably linked to a PLA2G1B transgene to direct expression of a PLA2G1B polypeptide to particular cells. A transgenic founder animal can be identified based upon the presence of a PLA2G1B transgene in its genome and/or expression of PLA2G1B mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover,

transgenic animals carrying a transgene encoding a PLA2G1B polypeptide can further be bred to other transgenic animals carrying other transgenes.

[0065] PLA2G1B polypeptides can be expressed in transgenic animals or plants by introducing, for example, a nucleic acid encoding the polypeptide into the genome of an animal. In preferred embodiments the nucleic acid is placed under the control of a tissue specific promoter, *e.g.*, a milk or egg specific promoter, and recovered from the milk or eggs produced by the animal. Also included is a population of cells from a transgenic animal.

PLA2G1B Polypeptides

[0066] Also featured herein are isolated PLA2G1B polypeptides, which include a polypeptide having the amino acid sequence of SEQ ID NO:2, PLA2G1B polypeptide variants, and substantially identical polypeptides thereof. A PLA2G1B polypeptide is a polypeptide encoded by a PLA2G1B nucleic acid, where one nucleic acid can encode one or more distinct polypeptides. An "isolated" or "purified" polypeptide or protein is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. In one embodiment, the language "substantially free" means preparation of a PLA2G1B polypeptide or PLA2G1B polypeptide variant having less than about 30%, 20%, 10% and more preferably 5% (by dry weight), of non-PLA2G1B polypeptide (also referred to herein as a "contaminating protein"), or of chemical precursors or non-PLA2G1B chemicals. When the PLA2G1B polypeptide or a biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, specifically, where culture medium represents less than about 20%, sometimes less than about 10%, and often less than about 5% of the volume of the polypeptide preparation. Isolated or purified PLA2G1B polypeptide preparations are sometimes 0.01 milligrams or more or 0.1 milligrams or more, and often 1.0 milligrams or more and 10 milligrams or more in dry weight.

[0067] Further included herein are PLA2G1B polypeptide fragments. The polypeptide fragment may be a domain or part of a domain of a PLA2G1B polypeptide. PLA2G1B domains include, but are not limited to, a phospholipase A2 domain at about amino acid positions 24 to 146 of SEQ ID NO:2. The polypeptide fragment may have increased, decreased or unexpected biological activity. The polypeptide fragment is often 50 or fewer, 100 or fewer, or 148 or fewer amino acids in length.

[0068] Substantially identical polypeptides may depart from the amino acid sequence of SEQ ID NO:2 in different manners. For example, conservative amino acid modifications may be introduced at one or more positions in the amino acid sequence of SEQ ID NO:2. A "conservative amino acid substitution" is one in which the amino acid is replaced by another amino acid having a similar structure and/or chemical function. Families of amino acid residues having similar structures and functions are

well known. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Also, essential and non-essential amino acids may be replaced. A "non-essential" amino acid is one that can be altered without abolishing or substantially altering the biological function of a PLA2G1B polypeptide, whereas altering an "essential" amino acid abolishes or substantially alters the biological function of a PLA2G1B polypeptide. Amino acids that are conserved among phospholipase A2 polypeptides (*e.g.*, P2X1, P2X2, P2X3, PLA2G1B, P2X5, P2X6, and P2X7) are typically essential amino acids.

[0069] Also, PLA2G1B polypeptides and polypeptide variants may exist as chimeric or fusion polypeptides. As used herein, a PLA2G1B "chimeric polypeptide" or "fusion polypeptide" includes a PLA2G1B polypeptide linked to a non-PLA2G1B polypeptide. A "non-PLA2G1B polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a polypeptide which is not substantially identical to the PLA2G1B polypeptide, which includes, for example, a polypeptide that is different from the PLA2G1B polypeptide and derived from the same or a different organism. The PLA2G1B polypeptide in the fusion polypeptide can correspond to an entire or nearly entire PLA2G1B polypeptide or a fragment thereof. The non-PLA2G1B polypeptide can be fused to the N-terminus or C-terminus of the PLA2G1B polypeptide.

[0070] Fusion polypeptides can include a moiety having high affinity for a ligand. For example, the fusion polypeptide can be a GST-PLA2G1B fusion polypeptide in which the PLA2G1B sequences are fused to the C-terminus of the GST sequences, or a polyhistidine-PLA2G1B fusion polypeptide in which the PLA2G1B polypeptide is fused at the N- or C-terminus to a string of histidine residues. Such fusion polypeptides can facilitate purification of recombinant PLA2G1B. Expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide), and a PLA2G1B nucleic acid can be cloned into an expression vector such that the fusion moiety is linked in-frame to the PLA2G1B polypeptide. Further, the fusion polypeptide can be a PLA2G1B polypeptide containing a heterologous signal sequence at its N-terminus. In certain host cells (*e.g.*, mammalian host cells), expression, secretion, cellular internalization, and cellular localization of a PLA2G1B polypeptide can be increased through use of a heterologous signal sequence. Fusion polypeptides can also include all or a part of a serum polypeptide (*e.g.*, an IgG constant region or human serum albumin).

[0071] PLA2G1B polypeptides can be incorporated into pharmaceutical compositions and administered to a subject *in vivo*. Administration of these PLA2G1B polypeptides can be used to affect the bioavailability of a PLA2G1B substrate and may effectively increase PLA2G1B biological activity in

a cell. PLA2G1B fusion polypeptides may be useful therapeutically for the treatment of disorders caused by, for example, (i) aberrant modification or mutation of a gene encoding a PLA2G1B polypeptide; (ii) mis-regulation of the PLA2G1B gene; and (iii) aberrant post-translational modification of a PLA2G1B polypeptide. Also, PLA2G1B polypeptides can be used as immunogens to produce anti-PLA2G1B antibodies in a subject, to purify PLA2G1B ligands or binding partners, and in screening assays to identify molecules which inhibit or enhance the interaction of PLA2G1B with a PLA2G1B substrate.

[0072] In addition, polypeptides of the invention can be chemically synthesized using techniques known in the art (See, e.g., Creighton, 1983 Proteins. New York, N.Y.: W. H. Freeman and Company; and Hunkapiller et al., (1984) Nature July 12 -18;310(5973):105-11). For example, a relative short fragment of the invention can be synthesized by use of a peptide synthesizer. Furthermore, if desired, nonclassical amino acids or chemical amino acid analogs can be introduced as a substitution or addition into the fragment sequence. Non-classical amino acids include, but are not limited to, to the D-isomers of the common amino acids, 2,4-diaminobutyric acid, a-amino isobutyric acid, 4-aminobutyric acid, Abu, 2-amino butyric acid, g-Abu, e-Ahx, 6-amino hexanoic acid, Aib, 2-amino isobutyric acid, 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, homocitrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine, b-alanine, fluoroamino acids, designer amino acids such as b-methyl amino acids, Ca-methyl amino acids, Na-methyl amino acids, and amino acid analogs in general. Furthermore, the amino acid can be D (dextrorotary) or L (levorotary).

[0073] The invention encompasses polypeptide fragments which are differentially modified during or after translation, e.g., by glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. Any of numerous chemical modifications may be carried out by known techniques, including but not limited, to specific chemical cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease, NaBH4; acetylation, formylation, oxidation, reduction; metabolic synthesis in the presence of tunicamycin; etc.

[0074] Additional post-translational modifications encompassed by the invention include, for example, N-linked or O-linked carbohydrate chains, processing of N-terminal or C-terminal ends), attachment of chemical moieties to the amino acid backbone, chemical modifications of N-linked or O-linked carbohydrate chains, and addition or deletion of an N-terminal methionine residue as a result of procaryotic host cell expression. The polypeptide fragments may also be modified with a detectable label, such as an enzymatic, fluorescent, isotopic or affinity label to allow for detection and isolation of the polypeptide.

[0075] Also provided are chemically modified derivatives of the polypeptides of the invention that may provide additional advantages such as increased solubility, stability and circulating time of the

polypeptide, or decreased immunogenicity. See U.S. Pat. No: 4,179,337. The chemical moieties for derivitization may be selected from water soluble polymers such as polyethylene glycol, ethylene glycol/propylene glycol copolymers, carboxymethylcellulose, dextran, polyvinyl alcohol and the like. The polypeptides may be modified at random positions within the molecule, or at predetermined positions within the molecule and may include one, two, three or more attached chemical moieties.

[0076] The polymer may be of any molecular weight, and may be branched or unbranched. For polyethylene glycol, the preferred molecular weight is between about 1 kDa and about 100 kDa (the term "about" indicating that in preparations of polyethylene glycol, some molecules will weigh more, some less, than the stated molecular weight) for ease in handling and manufacturing. Other sizes may be used, depending on the desired therapeutic profile (e.g., the duration of sustained release desired, the effects, if any on biological activity, the ease in handling, the degree or lack of antigenicity and other known effects of the polyethylene glycol to a therapeutic protein or analog).

[0077] The polyethylene glycol molecules (or other chemical moieties) should be attached to the polypeptide with consideration of effects on functional or antigenic domains of the polypeptide. There are a number of attachment methods available to those skilled in the art, e.g., EP 0 401 384, herein incorporated by reference (coupling PEG to G-CSF), see also Malik et al. (1992) Exp Hematol. September;20(8):1028-35, reporting pegylation of GM-CSF using tresyl chloride). For example, polyethylene glycol may be covalently bound through amino acid residues via a reactive group, such as, a free amino or carboxyl group. Reactive groups are those to which an activated polyethylene glycol molecule may be bound. The amino acid residues having a free amino group may include lysine residues and the N-terminal amino acid residues; those having a free carboxyl group may include aspartic acid residues, glutamic acid residues and the C-terminal amino acid residue. Sulfhydryl groups may also be used as a reactive group for attaching the polyethylene glycol molecules. Preferred for therapeutic purposes is attachment at an amino group, such as attachment at the N-terminus or lysine group.

[0078] One may specifically desire proteins chemically modified at the N-terminus. Using polyethylene glycol as an illustration of the present composition, one may select from a variety of polyethylene glycol molecules (by molecular weight, branching, etc.), the proportion of polyethylene glycol molecules to protein (polypeptide) molecules in the reaction mix, the type of pegylation reaction to be performed, and the method of obtaining the selected N-terminally pegylated protein. The method of obtaining the N-terminally pegylated preparation (i.e., separating this moiety from other monopegylated moieties if necessary) may be by purification of the N-terminally pegylated material from a population of pegylated protein molecules. Selective proteins chemically modified at the N-terminus may be accomplished by reductive alkylation, which exploits differential reactivity of different types of primary amino groups (lysine versus the N-terminal) available for derivatization in a particular protein. Under the

appropriate reaction conditions, substantially selective derivatization of the protein at the N-terminus with a carbonyl group containing polymer is achieved.

Substantially Identical PLA2G1B Nucleic Acids and Polypeptides

[0079] PLA2G1B nucleotide sequences and PLA2G1B polypeptide sequences that are substantially identical to the nucleotide sequence of SEQ ID NO:1 and the polypeptide sequence of SEQ ID NO:2. respectively, are included herein. The term "substantially identical" as used herein refers to two or more nucleic acids or polypeptides sharing one or more identical nucleotide sequences or polypeptide sequences, respectively. Included are nucleotide sequences or polypeptide sequences that are 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% (each often within a 1%, 2%, 3% or 4% variability) identical to the PLA2G1B nucleotide sequence in Figure 1 (SEQ ID NO:1) or the PLA2G1B polypeptide sequence of Figure 2 (SEQ ID NO:2). In certain embodiments, a nucleotide sequence substantially identical to the nucleotide sequence of SEQ ID NO:1 is 90% or more identical to the nucleotide sequence of SEQ ID NO:1 or encodes a polypeptide that is 90% or more identical to the polypeptide of SEQ ID NO:2. One test for determining whether two nucleic acids are substantially identical is to determine the percent of identical nucleotide sequences or polypeptide sequences shared between the nucleic acids or polypeptides.

[0080] Calculations of sequence identity are often performed as follows. Sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). The length of a reference sequence aligned for comparison purposes is sometimes 30% or more, 40% or more, 50% or more, often 60% or more, and more often 70%, 80%, 90%, 100% of the length of the reference sequence. The nucleotides or amino acids at corresponding nucleotide or polypeptide positions, respectively, are then compared among the two sequences. When a position in the first sequence is occupied by the same nucleotide or amino acid as the corresponding position in the second sequence, the nucleotides or amino acids are deemed to be identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, introduced for optimal alignment of the two sequences.

[0081] Comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. Percent identity between two amino acid or nucleotide sequences can be determined using the algorithm of E. Meyers and W. Miller, *CABIOS 4:* 11-17 (1989), which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4. Also, percent identity between two amino acid sequences can be determined using the Needleman and Wunsch, *J. Mol. Biol. 48:* 444-453 (1970)

algorithm which has been incorporated into the GAP program in the GCG software package (available at the http address www.gcg.com), using either a Blossum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. Percent identity between two nucleotide sequences can be determined using the GAP program in the GCG software package (available at http address www.gcg.com), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. A set of parameters often used is a Blossum 62 scoring matrix with a gap open penalty of 12, a gap extend penalty of 4, and a frameshift gap penalty of 5.

[0082] Another manner for determining if two nucleic acids are substantially identical is to assess whether a polynucleotide homologous to one nucleic acid will hybridize to the other nucleic acid under stringent conditions. As use herein, the term "stringent conditions" refers to conditions for hybridization and washing. Stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y., 6.3.1-6.3.6 (1989). Aqueous and non-aqueous methods are described in that reference and either can be used. An example of stringent hybridization conditions is hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 50°C. Another example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 55°C. A further example of stringent hybridization conditions is hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 60°C. Often, stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 65°C. More often, stringency conditions are 0.5M sodium phosphate, 7% SDS at 65°C, followed by one or more washes at 0.2X SSC, 1% SDS at 65°C.

[0083] An example of a substantially identical nucleotide sequence to SEQ ID NO:1 is one that has a different nucleotide sequence and still encodes the polypeptide sequence of SEQ ID NO:2. Another example is a nucleotide sequence that encodes a polypeptide having a polypeptide sequence that is more than 70% identical to, sometimes more than 75%, 80%, or 85% identical to, and often more than 90% and 95% identical to the polypeptide sequence of SEQ ID NO:2.

[0084] PLA2G1B nucleotide sequences and polypeptide sequences can be used as "query sequences" to perform a search against public databases to identify other family members or related sequences, for example. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul *et al., J. Mol. Biol. 215:* 403-10 (1990). BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to PLA2G1B nucleic acid molecules. BLAST polypeptide searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to

PLA2G1B polypeptides. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al., Nucleic Acids Res. 25(17):* 3389-3402 (1997). When utilizing BLAST and Gapped BLAST programs, default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used (*see* the http address www.ncbi.nlm.nih.gov).

[0085] A nucleic acid that is substantially identical to the nucleotide sequence of SEQ ID NO:1 may include polymorphic sites at positions equivalent to those described herein (*e.g.*, position 7328 in SEQ ID NO:1) when the sequences are aligned. For example, using the alignment procedures described herein, SNPs in a sequence substantially identical to the sequence of SEQ ID NO:1 can be identified at nucleotide positions that match (*i.e.*, align) with nucleotides at SNP positions in SEQ ID NO:1. Also, where a polymorphic variation is an insertion or deletion, insertion or deletion of a nucleotide sequence from a reference sequence can change the relative positions of other polymorphic sites in the nucleotide sequence.

[0086] Substantially identical PLA2G1B nucleotide and polypeptide sequences include those that are naturally occurring, such as allelic variants (same locus), splice variants, homologs (different locus), and orthologs (different organism) or can be non-naturally occurring. Non-naturally occurring variants can be generated by mutagenesis techniques, including those applied to polynucleotides, cells, or organisms. The variants can contain nucleotide substitutions, deletions, inversions and insertions. Variation can occur in either or both the coding and non-coding regions. The variations can produce both conservative and non-conservative amino acid substitutions (as compared in the encoded product). Orthologs, homologs, allelic variants, and splice variants can be identified using methods known in the art. These variants normally comprise a nucleotide sequence encoding a polypeptide that is 50%, about 55% or more, often about 70-75% or more, more often about 80-85% or more, and typically about 90-95% or more identical to the amino acid sequence shown in SEQ ID NO:2 or a fragment thereof. Such nucleic acid molecules can readily be identified as being able to hybridize under stringent conditions to the nucleotide sequence shown in SEQ ID NO:1 or a fragment of this sequence. Nucleic acid molecules corresponding to orthologs, homologs, and allelic variants of the PLA2G1B nucleotide sequence can further be identified by mapping the sequence to the same chromosome or locus as the PLA2G1B nucleotide sequence or variant.

[0087] Also, substantially identical PLA2G1B nucleotide sequences may include codons that are altered with respect to the naturally occurring sequence for enhancing expression of a PLA2G1B polypeptide or polypeptide variant in a particular expression system. For example, the nucleic acid can be one in which one or more codons are altered, and often 10% or more or 20% or more of the codons are altered for optimized expression in bacteria (*e.g.*, *E. coli.*), yeast (*e.g.*, *S. cervesiae*), human (*e.g.*, 293 cells), insect, or rodent (*e.g.*, hamster) cells.

Fat Deposition Disorder Prognostic and Diagnostic Methods

[0088] Methods for prognosing and diagnosing fat deposition, its related disorders (*e.g.*, obesity and NIDDM) and leanness in subjects are provided herein. These methods include detecting the presence or absence of one or more polymorphic variations in a PLA2G1B nucleotide sequence or substantially identical sequence thereof in a sample from a subject, where the presence of a polymorphic variant described herein is indicative of a predisposition to leanness or fat deposition or one or more fat deposition related disorders (*e.g.*, obesity or NIDDM). Determining a predisposition to fat deposition refers to determining whether an individual is at an increased or intermediate risk of fat deposition and determining a predisposition to leanness refers to a decreased risk of fat deposition. Determining a predisposition to NIDDM refers to determining whether an individual is at risk of NIDDM.

[0089] Thus, featured herein is a method for detecting a predisposition to fat deposition and a fat deposition disorder, such as obesity and NIDDM, in a subject, which comprises detecting the presence or absence of a polymorphic variation associated with fat deposition at a polymorphic site in a PLA2G1B nucleotide sequence in a nucleic acid sample from a subject, wherein the PLA2G1B nucleotide sequence comprises a polynucleotide sequence selected from the group consisting of: (a) the nucleotide sequence of SEQ ID NO:1; (b) a nucleotide sequence which encodes a polypeptide consisting of the amino acid sequence of SEQ ID NO:2; (c) a nucleotide sequence which encodes a polypeptide that is 90% identical to the amino acid sequence of SEQ ID NO:2 or a nucleotide sequence about 90% or more identical to the nucleotide sequence of SEQ ID NO:1; and (d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising the polymorphic site; whereby the presence of the polymorphic variation is indicative of a predisposition to fat deposition in the subject. In certain embodiments, polymorphic variants at positions 7328 and 9182 are detected for determining a predisposition to fat deposition, a polymorphic variant at position 7256 is detected for determining a predisposition to NIDDM and polymorphic variants at positions in linkage disequilibrium with these positions are detected for determining a predisposition to fat deposition and NIDDM. .

[0090] Results from prognostic tests may be combined with other test results to diagnose fat deposition related disorders, including NIDDM. For example, prognostic results may be gathered, a patient sample may be ordered based on a determined predisposition to fat deposition or NIDDM, the patient sample is analyzed, and the results of the analysis may be utilized to diagnose the fat deposition related condition (*e.g.*, NIDDM). Also fat deposition diagnostic methods can be developed from studies used to generate prognostic methods in which populations are stratified into subpopulations having different progressions of a fat deposition related disorder or condition.

[0091] Predisposition to fat deposition, fat deposition related disorders such as NIDDM and obesity, and leanness sometimes is expressed as a probability, such as an odds ratio, percentage, or risk factor. The predisposition is based upon the presence or absence of one or more polymorphic variants described

herein, and also may be based in part upon phenotypic traits of the individual being tested. Methods for calculating predispositions based upon patient data are well known (*see, e.g.*, Agresti, *Categorical Data Analysis*, 2nd Ed. 2002. Wiley). Allelotyping and genotyping analyses may be carried out in populations other than those exemplified herein to enhance the predictive power of the prognostic method. These further analyses are executed in view of the exemplified procedures described herein, and may be based upon the same polymorphic variations or additional polymorphic variations.

[0092] The nucleic acid sample typically is isolated from a biological sample obtained from a subject. For example, nucleic acid can be isolated from blood, saliva, sputum, urine, cell scrapings, and biopsy tissue. The nucleic acid sample can be isolated from a biological sample using standard techniques, such as the technique described in Example 2. As used herein, the term "subject" refers primarily to humans but also refers to other mammals such as dogs, cats, and ungulates (*e.g.*, cattle, sheep, and swine). Subjects also include avians (*e.g.*, chickens and turkeys), reptiles, and fish (*e.g.*, salmon), as embodiments described herein can be adapted to nucleic acid samples isolated from any of these organisms. The nucleic acid sample may be isolated from the subject and then directly utilized in a method for determining the presence of a polymorphic variant, or alternatively, the sample may be isolated and then stored (*e.g.*, frozen) for a period of time before being subjected to analysis.

[0093] The presence or absence of a polymorphic variant is determined using one or both chromosomal complements represented in the nucleic acid sample. Determining the presence or absence of a polymorphic variant in both chromosomal complements represented in a nucleic acid sample from a subject having a copy of each chromosome is useful for determining the zygosity of an individual for the polymorphic variant (*i.e.*, whether the individual is homozygous or heterozygous for the polymorphic variant). Any oligonucleotide-based diagnostic may be utilized to determine whether a sample includes the presence or absence of a polymorphic variant in a sample. For example, primer extension methods, ligase sequence determination methods (*e.g.*, U.S. Pat. Nos. 5,679,524 and 5,952,174, and WO 01/27326), mismatch sequence determination methods (*e.g.*, U.S. Pat. Nos. 5,851,770; 5,958,692; 6,110,684; and 6,183,958), microarray sequence determination methods, restriction fragment length polymorphism (RFLP), single strand conformation polymorphism detection (SSCP) (*e.g.*, U.S. Pat. Nos. 5,891,625 and 6,013,499), PCR-based assays (*e.g.*, TAQMAN® PCR System (Applied Biosystems)), and nucleotide sequencing methods may be used.

[0094] Oligonucleotide extension methods typically involve providing a pair of oligonucleotide primers in a polymerase chain reaction (PCR) or in other nucleic acid amplification methods for the purpose of amplifying a region from the nucleic acid sample that comprises the polymorphic variation. One oligonucleotide primer is complementary to a region 3' of the polymorphism and the other is complementary to a region 5' of the polymorphism. A PCR primer pair may be used in methods disclosed in U.S. Pat. Nos. 4,683,195; 4,683,202, 4,965,188; 5,656,493; 5,998,143; 6,140,054; WO

01/27327; and WO 01/27329 for example. PCR primer pairs may also be used in any commercially available machines that perform PCR, such as any of the GENEAMP® Systems available from Applied Biosystems. Also, those of ordinary skill in the art will be able to design oligonucleotide primers based upon the nucleotide sequence of SEQ ID NO:1 without undue experimentation using knowledge readily available in the art.

[0095] Also provided is an extension oligonucleotide that hybridizes to the amplified fragment adjacent to the polymorphic variation. As used herein, the term "adjacent" refers to the 3' end of the extension oligonucleotide being often 1 nucleotide from the 5' end of the polymorphic site, and sometimes 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from the 5' end of the polymorphic site, in the nucleic acid when the extension oligonucleotide is hybridized to the nucleic acid. The extension oligonucleotide then is extended by one or more nucleotides, and the number and/or type of nucleotides that are added to the extension oligonucleotide determine whether the polymorphic variant is present. Oligonucleotide extension methods are disclosed, for example, in U.S. Pat. Nos. 4,656,127; 4,851,331; 5,679,524; 5,834,189; 5,876,934; 5,908,755; 5,912,118; 5,976,802; 5,981,186; 6,004,744; 6,013,431; 6,017,702; 6,046,005; 6,087,095; 6,210,891; and WO 01/20039. Oligonucleotide extension methods using mass spectrometry are described, for example, in U.S. Pat. Nos. 5,547,835; 5,605,798; 5,691,141; 5,849,542; 5,869,242; 5,928,906; 6,043,031; and 6,194,144, and a method often utilized is described herein in Example 2.

[0096] A microarray can be utilized for determining whether a polymorphic variant is present or absent in a nucleic acid sample. A microarray may include any oligonucleotides described herein, and methods for making and using oligonucleotide microarrays suitable for diagnostic use are disclosed in U.S. Pat. Nos. 5,492,806; 5,525,464; 5,589,330; 5,695,940; 5,849,483; 6,018,041; 6,045,996; 6,136,541; 6,142,681; 6,156,501; 6,197,506; 6,223,127; 6,225,625; 6,229,911; 6,239,273; WO 00/52625; WO 01/25485; and WO 01/29259. The microarray typically comprises a solid support and the oligonucleotides may be linked to this solid support by covalent bonds or by non-covalent interactions. The oligonucleotides may also be linked to the solid support directly or by a spacer molecule. A microarray may comprise one or more oligonucleotides complementary to a polymorphic site of SEQ ID NO:1 (*e.g.*, positions 7256, 7328, and/or 9182).

[0097] A kit also may be utilized for determining whether a polymorphic variant is present or absent in a nucleic acid sample. A kit often comprises one or more pairs of oligonucleotide primers useful for amplifying a fragment of SEQ ID NO:1 or a substantially identical sequence thereof, where the fragment includes a polymorphic site. The kit sometimes comprises a polymerizing agent, for example, a thermostable nucleic acid polymerase such as one disclosed in U.S. Pat. Nos. 4,889,818 or 6,077,664. Also, the kit often comprises an elongation oligonucleotide that hybridizes to a PLA2G1B nucleic acid in a nucleic acid sample adjacent to the polymorphic site. Where the kit includes an elongation

oligonucleotide, it also often comprises chain elongating nucleotides, such as dATP, dTTP, dGTP, dCTP, and dITP, including analogs of dATP, dTTP, dGTP, dCTP and dITP, provided that such analogs are substrates for a thermostable nucleic acid polymerase and can be incorporated into a nucleic acid chain elongated from the extension oligonucleotide. Along with chain elongating nucleotides would be one or more chain terminating nucleotides such as ddATP, ddTTP, ddGTP, ddCTP, and the like. In an embodiment, the kit comprises one or more oligonucleotide primer pairs, a polymerizing agent, chain elongating nucleotides, at least one elongation oligonucleotide, and one or more chain terminating nucleotides. Kits optionally include buffers, vials, microtiter plates, and instructions for use.

[0098] Determining the presence of a polymorphic variant, or a combination of two or more polymorphic variants, in a PLA2G1B nucleic acid of the sample is often indicative of a predisposition to fat deposition, leanness, or NIDDM. For example, presence of a guanine at position 7328 of SEQ ID NO:1 in the sense strand of a PLA2G1B nucleotide sequence is associated with an increased risk of fat deposition and presence of an adenine at position 7328 of SEQ ID NO:1 in the sense strand of a PLA2G1B nucleotide sequence is associated with leanness or a decreased risk of fat deposition. Specifically, a subject homozygous for a guanine at position 7328 of SEQ ID NO:1 in the sense strands of the PLA2G1B nucleotide sequence is at a higher risk of fat deposition, a subject heterozygous for a guanine and adenine at position 7328 in the sense strands of the PLA2G1B nucleotide sequence is at an intermediate risk of increased fat deposition, and a subject homozygous for an adenine at position 7328 in the sense strands of the PLA2G1B nucleotide sequence is at a lower risk of fat deposition. Similarly, a subject homozygous for a cytosine at position 7328 in the strands complementary to the sense strands of the PLA2G1B nucleotide sequence is at a higher risk of increased fat deposition, a subject heterozygous for a cytosine and thymine at position 7328 in the strands complementary to the sense strands of the PLA2G1B nucleotide sequence is at an intermediate risk of increased fat deposition, and a subject homozygous for a thymine at position 7328 in the strands complementary to the sense strands of the PLA2G1B nucleotide sequence is at a decreased risk of fat deposition.

[0099] Also, presence of a thymine at position 9182 of SEQ ID NO:1 in the sense strand of a PLA2G1B nucleotide sequence is associated with an increased risk of fat deposition and the presence of a guanine at position 9182 in the sense strand of a PLA2G1B nucleotide sequence is associated with leanness or a decreased risk of fat deposition. Specifically, a subject homozygous for a thymine at position 9182 of SEQ ID NO:1 in the sense strands of the PLA2G1B nucleotide sequence is at a higher risk of increased fat deposition, a subject heterozygous for a thymine and guanine at position 9182 in the sense strands of the PLA2G1B nucleotide sequence is at an intermediate risk of increased fat deposition, and a subject homozygous for a guanine at position 9182 in the sense strands of the PLA2G1B nucleotide sequence is at a decreased risk of fat deposition. Similarly, a subject homozygous for an adenine at position 9182 in the strands complementary to the sense strands of the PLA2G1B nucleotide sequence is

at a higher risk of increased fat deposition, a subject heterozygous for an adenine and cytosine at position 9182 in the strands complementary to the sense strands of the PLA2G1B nucleotide sequence is at an intermediate risk of increased fat deposition, and a subject homozygous for a guanine at position 9182 in the strands complementary to the sense strands of the PLA2G1B nucleotide sequence is at a lower risk of fat deposition.

[00100] Also, the presence of a haplotypes of TTAG and GTAG at positions 4050, 7256, 7328, and 9182, respectively, in the sense strand of a PLA2G1B nucleotide sequence (SEQ ID NO:1) are associated with leanness or a decreased risk of fat deposition. Similarly, the presence of a haplotype of AATC and CATC at positions 4050, 7256, 7328, and 9182, respectively, in the strand complementary to the sense strand of a PLA2G1B nucleotide sequence are associated with leanness.

[00101] Presence of a cytosine at position 7256 of SEQ ID NO:1 in the sense strand of a PLA2G1B nucleotide sequence is associated with an increased risk of NIDDM and the presence of a thymine at position 7256 in the sense strand of a PLA2G1B nucleotide sequence is associated with a decreased risk of NIDDM. Specifically, a subject homozygous for a cytosine at position 7256 of SEQ ID NO:1 in the sense strands of the PLA2G1B nucleotide sequence is at a higher risk of NIDDM, a subject heterozygous for a cytosine and thymine at position 7256 in the sense strands of the PLA2G1B nucleotide sequence is at an intermediate risk of NIDDM, and a subject homozygous for a thymine at position 7256 in the sense strands of the PLA2G1B nucleotide sequence is at a decreased risk of NIDDM. Similarly, a subject homozygous for a guanine at position 7256 in the strands complementary to the sense strands of the PLA2G1B nucleotide sequence is at a higher risk of NIDDM, a subject heterozygous for an guanine and adenine at position 7256 in the strands complementary to the sense strands of the PLA2G1B nucleotide sequence is at an intermediate risk of NIDDM, and a subject homozygous for a adenine at position 7256 in the strands complementary to the sense strands of the PLA2G1B nucleotide sequence is at a lower risk of NIDDM.

Applications of Prognostic and Diagnostic Results to Pharmacogenomic Methods

[00102] Pharmacogenomics is a discipline that involves tailoring a treatment for a subject according to the subject's genotype as a particular treatment regimen may exert a differential effect depending upon the subject's genotype. Based upon the outcome of a prognostic test described herein, a clinician or physician may target pertinent information and preventative or therapeutic treatments to a subject who would be benefited by the information or treatment and avoid directing such information and treatments to a subject who would not be benefited (e.g., the treatment has no therapeutic effect and/or the subject experiences adverse side effects).

[00103] For example, where a candidate therapeutic exhibits a significant interaction with a major allele and a comparatively weak interaction with a minor allele (e.g., an order of magnitude or greater

difference in the interaction), such a therapeutic typically would not be administered to a subject genotyped as being homozygous for the minor allele, and sometimes not administered to a subject genotyped as being heterozygous for the minor allele. In another example, where a candidate therapeutic is not significantly toxic when administered to subjects who are homozygous for a major allele but is comparatively toxic when administered to subjects heterozygous or homozygous for a minor allele, the candidate therapeutic is not typically administered to subjects who are genotyped as being heterozygous or homozygous with respect to the minor allele.

[00104] The prognostic methods described herein are applicable to pharmacogenomic methods for preventing, alleviating or treating fat deposition conditions such as obesity and NIDDM. For example, a nucleic acid sample from an individual may be subjected to a prognostic test described herein. Where one or more polymorphic variations associated with increased risk of obesity or NIDDM are identified in a subject, information for preventing or treating obesity or NIDDM and/or one or more obesity or NIDDM treatment regimens then may be prescribed to that subject. For example, a patient having a cytosine at position 7256 in SEQ ID NO: 1 often is prescribed a preventative regimen designed to minimize the occurrence of NIDDM.

[00105] In certain embodiments, a treatment regimen is specifically prescribed and/or administered to individuals who will most benefit from it based upon their risk of developing obesity or NIDDM assessed by the prognostic methods described herein. Thus, provided are methods for identifying a subject predisposed to obesity or NIDDM and then prescribing a therapeutic or preventative regimen to individuals identified as having a predisposition. Thus, certain embodiments are directed to a method for reducing fat deposition, obesity or NIDDM in a subject, which comprises: detecting the presence or absence of a polymorphic variant associated with fat deposition, obesity or NIDDM in a PLA2G1B nucleotide sequence in a nucleic acid sample from a subject, where the PLA2G1B nucleotide sequence comprises a polynucleotide sequence selected from the group consisting of: (a) the polynucleotide sequence of SEQ ID NO:1; (b) a polynucleotide sequence which encodes a polypeptide consisting of the amino acid sequence of SEQ ID NO:2; (c) a polynucleotide sequence which encodes a polypeptide that is 90% identical to the amino acid sequence of SEQ ID NO:2; and (d) a fragment of a polynucleotide sequence of (a), (b), or (c); and prescribing or administering a treatment regimen to a subject from whom the sample originated where the presence of a polymorphic variation associated with fat reduction is detected in the PLA2G1B nucleotide sequence. In these methods, predisposition results may be utilized in combination with other test results to diagnose fat deposition associated conditions, such as obesity, metabolic conditions (e.g., NIDDM) and cardiovascular conditions (e.g., myocardial infarction).

[00106] The treatment sometimes is preventative (e.g., is prescribed or administered to reduce the probability that a fat deposition associated condition arises or progresses), sometimes is therapeutic, and sometimes delays, alleviates or halts the progression of a fat deposition associated condition. Any known

preventative or therapeutic treatment for alleviating or preventing the occurrence of a fat deposition associated disorder is prescribed and/or administered. For example, the treatment sometimes is or includes a drug that reduces fat deposition, including, for example, an appetite suppressant (e.g., Phentermine, Adipex, Bontril, Didrex, Ionamin, Meridia, Phendimetrazine, Tenuate, Sibutramine), a lipase inhibitor (e.g., Olistat), a phospholipase inhibitor, a PLA2G1B nucleic acid, a PLA2G1B polypeptide, and/or a molecule that interacts with a PLA2G1B nucleic acid or PLA2G1B polypeptide described hereafter. In another example, the treatment is or includes a physical exercise regimen, dietary counseling and/or a dietary regimen (e.g., a low fat diet and/or a diet where the subject eats during pre-scheduled intervals) optionally coupled with dietary counseling, psychological counseling and/or psychotherapy, and sometimes optionally coupled with prescription of a psychotherapeutic or psychoprophylactic (e.g., an antidepressant or anti-anxiety therapeutic). In other embodiments directed to diabetes management, a subject sometimes is prescribed a regimen for regularly monitoring blood glucose levels, dietary counseling, a dietary regimen for managing blood glucose levels, and/or a blood glucose altering drug regimen. Examples of blood glucose altering drug regimens are regular administration of insulin (e.g., injection, pump, inhaler spray, nasal spray, insulin patch, and insulin tablet), and administration of hypoglycemics (e.g., glyburide or repaglinide), starch blockers (e.g., acarbose), liver glucose regulating agents (e.g., metformin), and/or insulin sensitizers (e.g., rosiglitzaone or pioglitazone). Prescription and/or administration of each treatment or combinations of treatments often is dependent upon the age of the subject as well as the subject's physiological, medical, and/or psychological condition.

[00107] In an embodiment, the pharmacogenomic methods described herein are applicable to subjects who are women about forty or more years of age and have not yet entered menopause, undergoing menopause, or post-menopausal. Those subjects identified as having an increased risk for fat deposition sometimes are prescribed a hormone replacement treatment (HRT) regimen. There are many HRT regimens known in the art, which include regular administration of estrogen (e.g., Prumarin®), progesterone (e.g., Provera®), androgen (e.g., testosterone), a combination of estrogen and progesterone, a combination of estrogen and androgen (e.g., Estratest®), growth hormone, dehydroepiandrosterone (DHEA), a sulfate ester of DHEA, or a combination of DHEA and a DHEA sulfate ester. Also, selective estrogen receptor modulators (SERMs) such as raloxifene and tamoxifen, for example, can be prescribed. Those women diagnosed as having an increased risk of fat deposition sometimes are prescribed an estrogen replacement therapy (ERT) regimen or SERMs regimen as an alternative to a combination of estrogen and progesterone, due to an association between ERT and lower fat deposition and an association between increased fat deposition and progesterone replacement therapy.

[00108] In another embodiment, pharmacogenomic methods are applicable to subjects who are women using a contraceptive or are contemplating use of a contraceptive, where the contraceptive has

been shown to increase fat deposition in subjects. This embodiment often applies to women who are pre-pubescent, who are in puberty, or who are post-pubescent and pre-menopausal. Many oral contraceptives, especially those that include higher contents of estrogen compared to other oral contraceptives, have been shown to increase fat deposition in subjects. Those subjects identified as having an increased risk for fat deposition by the methods described herein often are advised not to begin an oral contraceptive regimen or to discontinue an oral contraceptive regimen. Alternatively, subjects identified as having an increased risk for fat deposition sometimes are advised to begin an oral contraceptive regimen using a contraceptive having lower estrogen content as compared to other available oral contraceptives (e.g., Allesse®, Levlite®, Loestrin-Fe®, and Mircette® are examples of contraceptives having lower estrogen content).

[00109] As therapeutic approaches for obesity or NIDDM continue to evolve and improve, the goal of treatments for fat deposition related disorders is to intervene even before clinical signs (e.g., impaired glucose tolerance) first manifest. Thus, genetic markers associated with susceptibility to obesity or NIDDM prove useful for early diagnosis, prevention and treatment of obesity or NIDDM.

[00110] As obesity or NIDDM preventative and treatment information can be specifically targeted to subjects in need thereof (e.g. those at risk of developing obesity or NIDDM or those that have early stages of obesity or NIDDM), provided herein is a method for preventing or reducing the risk of developing obesity or NIDDM in a subject, which comprises: (a) detecting the presence or absence of a polymorphic variation associated with obesity or NIDDM at a polymorphic site in a nucleotide sequence in a nucleic acid sample from a subject; (b) identifying a subject with a predisposition to obesity or NIDDM, whereby the presence of the polymorphic variation is indicative of a predisposition to obesity or NIDDM in the subject; and (c) if such a predisposition is identified, providing the subject with information about methods or products to prevent or reduce obesity or NIDDM or to delay the onset of obesity or NIDDM. Also provided is a method of targeting information or advertising to a subpopulation of a human population based on the subpopulation being genetically predisposed to a disease or condition, which comprises: (a) detecting the presence or absence of a polymorphic variation associated with obesity or NIDDM at a polymorphic site in a nucleotide sequence in a nucleic acid sample from a subject; (b) identifying the subpopulation of subjects in which the polymorphic variation is associated with obesity or NIDDM; and (c) providing information only to the subpopulation of subjects about a particular product which may be obtained and consumed or applied by the subject to help prevent or delay onset of the disease or condition.

[00111] Pharmacogenomics methods also may be used to analyze and predict a response to an obesity or NIDDM treatment or a drug. For example, if pharmacogenomics analysis indicates a likelihood that an individual will respond positively to a obesity or NIDDM treatment with a particular drug, the drug may be administered to the individual. Conversely, if the analysis indicates that an individual is likely to

respond negatively to treatment with a particular drug, an alternative course of treatment may be prescribed. A negative response may be defined as either the absence of an efficacious response or the presence of toxic side effects. The response to a therapeutic treatment can be predicted in a background study in which subjects in any of the following populations are genotyped: a population that responds favorably to a treatment regimen, a population that does not respond significantly to a treatment regimen, and a population that responds adversely to a treatment regiment (e.g. exhibits one or more side effects). These populations are provided as examples and other populations and subpopulations may be analyzed. Based upon the results of these analyses, a subject is genotyped to predict whether he or she will respond favorably to a treatment regimen, not respond significantly to a treatment regimen, or respond adversely to a treatment regimen.

[00112] The prognostic tests described herein also are applicable to clinical drug trials. One or more polymorphic variants indicative of response to an agent for treating obesity or NIDDM or to side effects to an agent for treating obesity or NIDDM may be identified using the methods described herein. Thereafter, potential participants in clinical trials of such an agent may be screened to identify those individuals most likely to respond favorably to the drug and exclude those likely to experience side effects. In that way, the effectiveness of drug treatment may be measured in individuals who respond positively to the drug, without lowering the measurement as a result of the inclusion of individuals who are unlikely to respond positively in the study and without risking undesirable safety problems.

[00113] Thus, another embodiment is a method of selecting an individual for inclusion in a clinical trial of a treatment or drug comprising the steps of: (a) obtaining a nucleic acid sample from an individual; (b) determining the identity of a polymorphic variation which is associated with a positive response to the treatment or the drug, or at least one polymorphic variation which is associated with a negative response to the treatment or the drug in the nucleic acid sample, and (c) including the individual in the clinical trial if the nucleic acid sample contains said polymorphic variation associated with a positive response to the treatment or the drug or if the nucleic acid sample lacks said polymorphic variation associated with a negative response to the treatment or the drug. In addition, the methods of the present invention for selecting an individual for inclusion in a clinical trial of a treatment or drug encompass methods with any further limitation described in this disclosure, or those following, specified alone or in any combination. The polymorphic variation may be in a sequence selected individually or in any combination from the group consisting of (i) a polynucleotide sequence set forth in SEQ ID NO: 1; (ii) a polynucleotide sequence that is 90% identical to a nucleotide sequence set forth in SEQ ID NO: 1; (iii) a polynucleotide sequence that encodes a polypeptide having an amino acid sequence identical to or 90% identical to an amino acid sequence encoded by a nucleotide sequence set forth in SEQ ID NO: 1; and (iv) a fragment of a polynucleotide sequence of (i), (ii), or (iii) comprising the polymorphic site. The including step (c) optionally comprises administering the drug or the treatment to the individual if the

nucleic acid sample contains the polymorphic variation associated with a positive response to the treatment or the drug and the nucleic acid sample lacks said biallelic marker associated with a negative response to the treatment or the drug.

[00114] Also provided herein is a method of partnering between a diagnostic/prognostic testing provider and a provider of a consumable product, which comprises: (a) the diagnostic/prognostic testing provider detects the presence or absence of a polymorphic variation associated with obesity or NIDDM at a polymorphic site in a nucleotide sequence in a nucleic acid sample from a subject; (b) the diagnostic/prognostic testing provider identifies the subpopulation of subjects in which the polymorphic variation is associated with obesity or NIDDM; (c) the diagnostic/prognostic testing provider forwards information to the subpopulation of subjects about a particular product which may be obtained and consumed or applied by the subject to help prevent or delay onset of the disease or condition; and (d) the provider of a consumable product forwards to the diagnostic test provider a fee every time the diagnostic/prognostic test provider forwards information to the subject as set forth in step (c) above.

Methods for Identifying Candidate Therapeutics for Reducing Fat Deposition and Treating Related Disorders

[00115] Current therapies for the treatment of NIDDM have limited efficacy, limited tolerability and significant mechanisms-based side effects, including weight gain and hypoglycaemia. Few of the available therapies adequately address underlying defects such as obesity and insulin resistance. Thus, newer approaches are desperately needed (Moller D. Nature. 414:821-827 (2001)). Current therapeutic approaches were largely developed in the absence of defined molecular targets or even a solid understanding of disease pathogenesis. The same holds true for the treatment of obesity, where treatments have limited lasting effects and many side effects. Therefore, there is a need for methods of identifying candidate therapeutics that target the biochemical pathways related to the development of obesity and/or diabetes.

[00116] Featured herein is a method for identifying candidate therapeutics for reducing fat deposition and/or the development of NIDDM. The method comprises contacting a test molecule with a PLA2G1B nucleic acid, nucleic acid variant, polypeptide, or polypeptide variant in a system. The nucleic acid is often the PLA2G1B nucleotide sequence represented by SEQ ID NO:1, sometimes a nucleotide sequence that is substantially identical to the nucleotide sequence of SEQ ID NO:1, or sometimes a fragment thereof, and the PLA2G1B polypeptide is a polypeptide encoded by any of these nucleic acids. The method also comprises determining the presence or absence of an interaction between the test molecule and the PLA2G1B nucleic acid or polypeptide, where the presence of an interaction between the test molecule and the PLA2G1B nucleic acid or polypeptide identifies the test molecule as a candidate therapeutic for fat reduction or NIDDM.

[00117] As discussed, successful treatment of PLA2G1B disorders can be brought about by techniques that serve to inhibit the expression or activity of target gene products. For example, compounds (*e.g.*, an agent identified using an assays described above or an siRNA molecule) that exhibit negative modulatory activity can be used in accordance with the invention to prevent and/or ameliorate fat deposition or diabetes. Such molecules can include, but are not limited to peptides, phosphopeptides, small organic or inorganic molecules, or antibodies (including, for example, polyclonal, monoclonal, humanized, anti-idiotypic, chimeric or single chain antibodies, and FAb, $F(ab')_2$ and FAb expression library fragments, scFV molecules, and epitope-binding fragments thereof).

[00118] Further, antisense and ribozyme molecules that inhibit expression of the target gene can also be used in accordance with the invention to reduce the level of target gene expression, thus effectively reducing the level of target gene activity. Still further, triple helix molecules can be utilized in reducing the level of target gene activity. Antisense, ribozyme and triple helix molecules are discussed above.

[00119] It is possible that the use of antisense, ribozyme, and/or triple helix molecules to reduce or inhibit mutant gene expression can also reduce or inhibit the transcription (triple helix) and/or translation (antisense, ribozyme) of mRNA produced by normal target gene alleles, such that the concentration of normal target gene product present can be lower than is necessary for a normal phenotype. In such cases, nucleic acid molecules that encode and express target gene polypeptides exhibiting normal target gene activity can be introduced into cells via gene therapy method. Alternatively, in instances in that the target gene encodes an extracellular polypeptide, it can be preferable to co-administer normal target gene polypeptide into the cell or tissue in order to maintain the requisite level of cellular or tissue target gene activity.

[00120] PLA2G1B gene expression sometimes can be inhibited by the introduction of double-stranded RNA (dsRNA), which induces potent and specific gene silencing, a phenomenon called RNA interference or RNAi. See, *e.g.*, Fire *et al.*, US Patent Number 6,506,559; Tuschl *et al.* PCT International Publication No. WO 01/75164; Kay *et al.* PCT International Publication No. WO 03/010180A1; or Bosher JM, Labouesse, *Nat Cell Biol* 2000 Feb;2(2):E31-6. This process has been improved by decreasing the size of the double-stranded RNA to 20-24 base pairs (to create small-interfering RNAs or siRNAs) that "switched off" genes in mammalian cells without initiating an acute phase response, *i.e.*, a host defense mechanism that often results in cell death. *See, e.g.*, Caplen et al. *Proc Natl Acad Sci* U S A. 2001 Aug 14;98(17):9742-7 and Elbashir SM et al. *Methods* 2002 Feb;26(2):199-213.

[00121] There is increasing evidence that post-transcriptional gene silencing by RNA interference (RNAi) for inhibiting targeted expression in mammalian cells at the mRNA level is effective in human cells. There is additional evidence of effective methods for inhibiting the proliferation and migration of tumor cells in human patients, and for inhibiting metastatic cancer development. *See, e.g.*, U.S. Patent

Application Number US2001000993183; Caplen NJ et al. *Proc Natl Acad Sci* U S A; and Abderrahmani A. et al. *Mol Cell Biol* 2001 Nov21(21):7256-67.

[00122] An "siRNA" or "RNAi" refers to a nucleic acid that forms a double stranded RNA and has the ability to reduce or inhibit expression of a gene or target gene when the siRNA is delivered to or expressed in the same cell as the gene or target gene. "siRNA" thus refers to short double stranded RNA formed by the complementary strands. Complementary portions of the siRNA that hybridize to form the double stranded molecule often have substantial or complete identity to the target molecule sequence. In one embodiment, an siRNA refers to a nucleic acid that has substantial or complete identity to a target gene and forms a double stranded siRNA, such as a nucleotide sequence in SEQ ID NO: 1, for example.

[00123] When designing the siRNA molecules, the targeted region often is selected from a given DNA sequence beginning 50 to 100 nt downstream of the start codon. *See, e.g.*, Elbashir *et al,. Methods 26*:199-213 (2002). Initially, 5' or 3' UTRs and regions nearby the start codon were avoided assuming that UTR-binding proteins and/or translation initiation complexes may interfere with binding of the siRNP or RISC endonuclease complex. Sometimes regions of the target 23 nucleotides in length conforming to the sequence motif AA(N19)TT (N, an nucleotide), and regions with approximately 30% to 70% G/C-content (often about 50% G/C-content) often are selected. If no suitable sequences are found, the search often is extended using the motif NA(N21). The sequence of the sense siRNA sometimes corresponds to (N19) TT or N21 (position 3 to 23 of the 23-nt motif), respectively. In the latter case, the 3' end of the sense siRNA often is converted to TT. The rationale for this sequence conversion is to generate a symmetric duplex with respect to the sequence composition of the sense and antisense 3' overhangs. The antisense siRNA is synthesized as the complement to position 1 to 21 of the 23-nt motif. Because position 1 of the 23-nt motif is not recognized sequence-specifically by the antisense siRNA, the 3'-most nucleotide residue of the antisense siRNA can be chosen deliberately. However, the penultimate nucleotide of the antisense siRNA (complementary to position 2 of the 23-nt motif) often is complementary to the targeted sequence. For simplifying chemical synthesis, TT often is utilized. siRNAs corresponding to the target motif NAR(N17)YNN, where R is purine (A,G) and Y is pyrimidine (C,U), often are selected. Respective 21 nucleotide sense and antisense siRNAs often begin with a purine nucleotide and can also be expressed from pol III expression vectors without a change in targeting site. Expression of RNAs from pol III promoters often is efficient when the first transcribed nucleotide is a purine.

[00124] The sequence of the siRNA can correspond to the full length target gene, or a subsequence thereof. Often, the siRNA is about 15 to about 50 nucleotides in length (*e.g.*, each complementary sequence of the double stranded siRNA is 15-50 nucleotides in length, and the double stranded siRNA is about 15-50 base pairs in length, somtimes about 20-30 nucleotides in length or about 20-25 nucleotides in length, *e.g.*, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleotides in length. The siRNA often is

about 21 nucleotides in length. Methods of using siRNA are well known in the art, and specific siRNA molecules may be purchased from a number of companies including Dharmacon Research, Inc.

[00125] Another method by which nucleic acid molecules may be utilized in treating or preventing a disease characterized by PLA2G1B expression is through the use of aptamer molecules specific for PLA2G1B polypeptide. Aptamers are nucleic acid molecules having a tertiary structure which permits them to specifically bind to polypeptide ligands (see, *e.g.*, Osborne, *et al., Curr. Opin. Chem. Biol.1(1):* 5-9 (1997); and Patel, D. J., *Curr. Opin. Chem. Biol. Jun;1(1):* 32-46 (1997)). Since nucleic acid molecules may in many cases be more conveniently introduced into target cells than therapeutic polypeptide molecules may be, aptamers offer a method by which PLA2G1B polypeptide activity may be specifically decreased without the introduction of drugs or other molecules which may have pluripotent effects.

[00126] Antibodies can be generated that are both specific for target gene product and that reduce target gene product activity. Such antibodies may, therefore, by administered in instances whereby negative modulatory techniques are appropriate for the treatment of PLA2G1B disorders. For a description of antibodies, see the Antibody section above.

[00127] In circumstances where injection of an animal or a human subject with a PLA2G1B polypeptide or epitope for stimulating antibody production is harmful to the subject, it is possible to generate an immune response against PLA2G1B through the use of anti-idiotypic antibodies (see, for example, Herlyn, D., *Ann. Med.;31(1):* 66-78 (1999); and Bhattacharya-Chatterjee, M., and Foon, K.A., *Cancer Treat. Res.; 94:* 51-68 (1998)). If an anti-idiotypic antibody is introduced into a mammal or human subject, it should stimulate the production of anti-anti-idiotypic antibodies, which should be specific to the PLA2G1B polypeptide. Vaccines directed to a disease characterized by PLA2G1B expression may also be generated in this fashion.

[00128] In instances where the target antigen is intracellular and whole antibodies are used, internalizing antibodies may be preferred. Lipofectin or liposomes can be used to deliver the antibody or a fragment of the Fab region that binds to the target antigen into cells. Where fragments of the antibody are used, the smallest inhibitory fragment that binds to the target antigen is preferred. For example, peptides having an amino acid sequence corresponding to the Fv region of the antibody can be used. Alternatively, single chain neutralizing antibodies that bind to intracellular target antigens can also be administered. Such single chain antibodies can be administered, for example, by expressing nucleotide sequences encoding single-chain antibodies within the target cell population (see *e.g.*, Marasco *et al., Proc. Natl. Acad. Sci. USA 90:* 7889-7893 (1993)).

[00129] PLA2G1B molecules and compounds that inhibit target gene expression, synthesis and/or activity can be administered to a patient at therapeutically effective doses to prevent, treat or ameliorate

PLA2G1B disorders. A therapeutically effective dose refers to that amount of the compound sufficient to result in amelioration of symptoms of the disorders.

[00130] Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD_{50}/ED_{50}. Compounds that exhibit large therapeutic indices are preferred. While compounds that exhibit toxic side effects can be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

[00131] Data obtained from cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage can vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose can be formulated in animal models to achieve a circulating plasma concentration range that includes the IC_{50} (*i.e.*, the concentration of the test compound that achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma can be measured, for example, by high performance liquid chromatography.

[00132] Another example of effective dose determination for an individual is the ability to directly assay levels of "free" and "bound" compound in the serum of the test subject. Such assays may utilize antibody mimics and/or "biosensors" that have been created through molecular imprinting techniques. The compound which is able to modulate PLA2G1B activity is used as a template, or "imprinting molecule", to spatially organize polymerizable monomers prior to their polymerization with catalytic reagents. The subsequent removal of the imprinted molecule leaves a polymer matrix which contains a repeated "negative image" of the compound and is able to selectively rebind the molecule under biological assay conditions. A detailed review of this technique can be seen in Ansell, R. J. *et al., Current Opinion in Biotechnology 7:* 89-94 (1996) and in Shea, K.J., *Trends in Polymer Science 2:* 166-173 (1994). Such "imprinted" affinity matrixes are amenable to ligand-binding assays, whereby the immobilized monoclonal antibody component is replaced by an appropriately imprinted matrix. An example of the use of such matrixes in this way can be seen in Vlatakis, G. *et al., Nature 361:* 645-647 (1993). Through the use of isotope-labeling, the "free" concentration of compound which modulates the expression or activity of PLA2G1B can be readily monitored and used in calculations of IC_{50}. Such "imprinted" affinity matrixes can also be designed to include fluorescent groups whose photon-emitting

properties measurably change upon local and selective binding of target compound. These changes can be readily assayed in real time using appropriate fiberoptic devices, in turn allowing the dose in a test subject to be quickly optimized based on its individual IC_{50}. A rudimentary example of such a "biosensor" is discussed in Kriz, D. *et al., Analytical Chemistry 67:* 2142-2144 (1995).

[00133] Provided herein are methods of modulating PLA2G1B expression or activity for therapeutic purposes. Accordingly, in an exemplary embodiment, the modulatory method of the invention involves contacting a cell with a PLA2G1B or agent that modulates one or more of the activities of PLA2G1B polypeptide activity associated with the cell. An agent that modulates PLA2G1B polypeptide activity can be an agent as described herein, such as a nucleic acid or a polypeptide, a naturally-occurring target molecule of a PLA2G1B polypeptide (*e.g.*, a PLA2G1B substrate or receptor), a PLA2G1B antibody, a PLA2G1B agonist or antagonist, a peptidomimetic of a PLA2G1B agonist or antagonist, or other small molecule.

[00134] In one embodiment, the agent stimulates one or more PLA2G1B activities. Examples of such stimulatory agents include active PLA2G1B polypeptide and a nucleic acid molecule encoding PLA2G1B. In another embodiment, the agent inhibits one or more PLA2G1B activities. Examples of such inhibitory agents include antisense PLA2G1B nucleic acid molecules, anti-PLA2G1B antibodies, and PLA2G1B inhibitors. These modulatory methods can be performed *in vitro* (*e.g.*, by culturing the cell with the agent) or, alternatively, *in vivo* (*e.g.*, by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant or unwanted expression or activity of a PLA2G1B polypeptide or nucleic acid molecule. In one embodiment, the method involves administering an agent (*e.g.*, an agent identified by a screening assay described herein), or combination of agents that modulates (*e.g.*, upregulates or downregulates) PLA2G1B expression or activity. In another embodiment, the method involves administering a PLA2G1B polypeptide or nucleic acid molecule as therapy to compensate for reduced, aberrant, or unwanted PLA2G1B expression or activity.

[00135] The examples set forth below are intended to illustrate but not limit the invention.

Examples

[00136] In the following studies a group of subjects were selected according to specific parameters relating to fat deposition. Nucleic acid samples obtained from individuals in the study group were subjected to genetic analysis, which identified associations between central obesity and certain polymorphic regions in the PLA2G1B gene on chromosome 12. Polymorphic variations identified as being associated with central obesity were further screened in subjects with NIDDM to determine if they are also associated with the development of diabetes. Methods are described for producing PLA2G1B polypeptide and PLA2G1B polypeptide variants *in vitro* or *in vivo*, PLA2G1B nucleic acids or

polypeptides and variants thereof are utilized for screening test molecules for those that interact with PLA2G1B molecules. Test molecules identified as interactors with PLA2G1B molecules and PLA2G1B variants are further screened *in vivo* to determine whether they can reduce fat deposition or treat diabetes.

Example 1

Sample Selection

[00137] In addition to simple clinical measurements, dual x-ray absorbtiometry (DEXA) was utilized to determine fat content in subjects for the genetic analysis. Central fat was the primary target variable, and data were collected using a Hologic QDR 4500 DEXA system. The central region for central fat determinations was defined as the region extending from the superior surface of the second lumbar vertebra extending inferiorly to the inferior surface of the fourth lumbar vertebra and laterally to the inner aspect of the ribcage. The amount of central fat and percent central fat was automatically calculated by the equipment and downloaded into a database.

[00138] Waist and hip measurements were generated while subjects were wearing underclothes and standing with their arms by their sides. A tape measure was utilized for these measurements, and care was taken to ensure that the tape was resting on the skin and not tight. Waist circumference was measured to the nearest centimeter at the narrowest point between the iliac crest and the lower edge of the ribs. Hip circumference was measured to the nearest centimeter at the widest point below the iliac crest.

[00139] Sample selection was restricted to female twins followed by the St. Thomas Hospital in England and the Royal North Shore Hospital in Australia. It was estimated that 552 dizygotic sibling pairs would yield statistical results of reasonable power. A further 272 unrelated individuals selected from monozygotic twin pairs were added to the sample set to increase the probability for detecting associations and also for testing gene-environment interactions. The study group was selected from this combination of dizygotic and monozygotic sibling pairs, referred to as the "selection group."

[00140] Central fat measurements and triglyceride measurements were chosen as primary target phenotypes and twin pairs were selected from the selection group for extreme discordance and concordance. Specifically, DEXA measurements and triglyceride measurements (colorimetric enzymatic method: glycerol-3-phosphate-oxidase, peroxidase, PAP (Roche), CV% = 2.6, reference range less than 2.5 mmol/L) for each individual in the selection group were arranged in ascending order, and individuals in the top and lower tenth percentile were chosen from each distribution. A small subset of individuals falling in the middle range of each distribution was chosen as a control group. In addition to primary phenotype trait information, samples for inclusion in the study group were selected based on data coverage for the following secondary phenotypes recorded by each individual: BMI, insulin resistance, high density lipoprotein in serum, waist, lipoprotein(a) in serum, insulin, hip, and waist/hip ratio.

[00141] Also, presence of diabetes, thyroid disease, and renal disease reported by each individual were primary criteria for excluding subjects from the study group. Also, insulin levels greater than 7.1μU/ml (Microparticle Enzyme Immunoassay from Abbott Laboratories Diagnostics Division (μU/ml = pmol/L x 7.175) and creatine levels greater than 160 μmol/L (measured by Jaffe method: calorimetric test in which creatine reacts with picric acid in an alkaline solution to form a yellow-red colored complex) were also used as exclusion criteria as they are indicative of these diseases. Further excluded were pairs discordant for menopausal status, twin pairs where one or both of the twins were taking lipid lowering medication, non-fasting subjects (less than eight hours eating), and twin pairs including subjects treated with beta-blockers, thiazide diuretics, or exogenous estrogen.

[00142] Selecting among dizygotic and monozygotic twins for extreme discordance or concordance for the primary phenotypes minimized complications associated with bivariate ranking. After applying exclusion criteria, 253 monozygotic subjects were available for inclusion, which fell short of the target population of 276. In reaction to this situation, the extreme 201 subjects were selected from the 253 subjects and the desired numbers were reached by adding monozygotic unrelated individuals with data for central fat only, and unrelated individuals from the dizygotic cohort with data for triglycerides only. Samples available for final selection for the 552 dizygotic pairs included 178 pairs extreme for both traits, 205 extreme for triglycerides only, and 208 for central fat only.

[00143] In this test population, coverage for the secondary phenotypes ranged from 67% to 90%. In total, 61% of subjects had coverage for all primary and secondary phenotypes. A broad age spectrum was also represented, and numbers of pre-menopausal and post-menopausal subjects were relatively evenly distributed.

Example 2

Association of Polymorphic Variations to Fat Deposition

[00144] Blood samples were taken from individuals in the study population described in Example 1. Genomic DNA was extracted from these blood samples using standard techniques (BACC2 DNA extraction kit (Nucleon Biosciences)) and subjected to analysis. Based upon a background linkage study and fine mapping analysis by microsatellite markers, it was postulated that genetic elements linked to central fat deposition were located on the 12q24 region of chromosome twelve. One of the genes located in this region encoded PLA2G1B.

[00145] Whole genome linkage scans were performed for the purpose of identifying genomic regions likely to harbor genes with a major contribution to deposition of central fat. The linkage scans were performed using highly polymorphic microsatellite markers (Reed *et al., Nat Genet 7*:390-5 (1994)) and DNA samples obtained from 1100 Caucasian female twin pairs from the UK. Samples selected for inclusion in the study cohort encompassed a broad spectrum of phenotypic trait values, ranging from lean

to obese subjects. Initial studies were carried out using 400 commercially available microsatellite markers derived from the Genethon linkage map, with an average genomic spacing between markers of approximately 10 cM (ABI Prism linkage mapping set, version 2 from PE Applied Biosystems).

[00146] Multipoint nonparametric linkage analysis was performed using MAPMAKER/SIBS (Kruglyak & Lander, *Amer. J. Human Genetics 57*:439-454 (1995)). A bioinformatics infrastructure and software packages described in WO 00/51053 were used in the linkage study to record marker positions, store data and generate data files. Output from these systems was then used with relevant application software to perform the statistical analysis.

[00147] Genotyping reactions were generally carried out in microtitre plates (384-well, reaction volume 5μl), containing 12.5ng of DNA from study subjects was amplified using PCR and sequence specific oligonucleotide primers labeled with 6-FAM™, HEX™, or NED™ fluorescent dyes. PCR products were analyzed by electrophoresis in a polyacrylamide denaturing gel, with an ABI PRISM™ GENESCAN® 400HD ROX labeled size standard in each lane on an ABI model 377 analyzer (Applied Biosystems, Foster City, California). For genotyping, the chosen markers were divided into two groups (panels) so that the analysis of all of the markers could be performed in two electrophoresis runs of each sample. Consequently, there was no overlap of fragment sizes in any one dye for either of the panels. Genotype analysis was performed using ABI PRISM™ GENESCAN® software (version 3.0), and genotyped manually using ABI PRISM™ Genotyper 2.0. Results were input into a database and binned by marker. The results were quality checked, ensuring consistent inheritance within families. Families that were found to have consistent pedigree problems were excluded from the analysis set.

[00148] The ordering of genetic mapping markers (*i.e.* microsatellite markers) was relatively stable in the region analyzed according to the Unified Data Base for Human Genome Mapping, Weizmann Institute of Science (UDB) and National Center for Biotechnology Information, National Institutes of Health (NCBI) assemblies during the duration of the study. Conversion of genetic to physical positions for strategic microsatellite markers was performed using UDB and NCBI as the reference standards. Comparisons of the identity and positioning of genomic contigs in the region were also made between UDB and NCBI and provided relatively good agreement. A comparison of the positioning of all identified and predicted genes within the region was also made between NCBI (build 22) and Joint Project between European Bioinformatics Institute and the Sanger Centre (ENSEMBL).

[00149] Microsatellite marker analysis showed linkage on the long arm of chromosome 12 to central fat deposition, percent central fat and total fat in the region spanning 125 cM to 155 cM, with a peak non parametric Z score of 3.6 for central fat. The region was further narrowed to identify the chromosomal interval 12q24 as being the primary region harboring genes contributing to central fat deposition using the following highly polymorphic microsatellite markers: D12S86, D12S1612, D12S1614, D12S340, D12S324, D12S1675, D12S1679, D12S1659 and D12S97.

[00150] The chromosome 12q24 region was then analyzed using single nucleotide polymorphisms to identify genes in the region that regulate central fat deposition. Potential polymorphisms in the PLA2G1B polynucleotide were identified in a publicly available SNP database (*see* http address www.ncbi.nlm.nih.gov/SNP) and were verified in a group other than the study group. Polymorphisms verified as statistically significant SNPs (minor allele represented in more than 10% of the population) were genotyped in the study population to determine associations with fat deposition. A procedure for detecting polymorphisms was utilized in the verification and genotyping studies, described hereafter. Table 1 shows the majority of polymorphisms subjected to genotype analysis and allelic variability reported in dbSNP.

Table 1

Reference SNP ID	Position in SEQ ID NO:1	Reported Allelic Variability
rs2701632	436	T/C
rs2009391	4050	A/C
rs5631	4689	T/A
rs5632	6282	A/C
rs5633	6358	C/T
rs5634	7256	T/C
rs5635	7300	A/C
rs5636	7301	C/A
rs5637	7328	G/A
rs1186217	8062	C/T
rs1179387	9182	T/G
rs2701629	11649	C/A
rs2701631	839	A/T
rs2070873	6653	T/G
rs2066539	10164	G/A

Assays for Verifying and Genotyping SNPs

[00151] An assay utilized for determining whether a polymorphic variation was present in a nucleic acid sample involved a sequencing by synthesis procedure. DNA polymerase, ATP sulfurylase, luciferase, apyrase, luciferin, and adenosine 5'-phosphosulfate (APS) were utilized, and in the process, one dNTP was added to an extension oligonucleotide at a time and then degraded if not incorporated in the synthesized strand. Incorporation of a dNTP to the end of the extension oligonucleotide was detected by light emission.

[00152] The assay was carried out by first amplifying a region of interest in the sample by using a polymerase chain reaction (PCR) that incorporated the primers set forth in Table 2.

Table 2

Reference SNP ID	Position in SEQ ID NO:1	First PCR primer	SEQ ID NO:	Second PCR primer	SEQ ID NO:
rs2009391	4050	TGCAGAGGCTCAATCA CTGT		CAGGTGTGGTGGTGGA TTG	
rs5631	4689	CACAGGCCACAGCAA ACAG		TCAGACTTGCAGGTTGA AAAAG	
rs5632	6282	GGCAGACCGATTTGAA CTCT		CGGGATCACGCACTTG A	
rs5633	6358	GGCAGTTCCGCAAAAT GAT		TGCAGGCGGATCACTT ACTT	
rs5634	7256	AGCTGTCCCTCCCACT TTC		GTGTGGGTGTACGGGTT GT	
rs5635	7300	AGCTGTCCCTCCCACT TTC		ATAGGTCAAGGAAGGG ATAAAC	
rs5636	7301	AGCTGTCCCTCCCACT TTC		ATAGGTCAAGGAAGGG ATAAAC	
rs5637	7328	CAAGAAGCTGGACAG CTGTA		ATAGGTCAAGGAAGGG ATAAAC	
rs1186217	8062	ATCACCTCAACCTCCG TTCA		GGTGGTGCACGCTTGTA ATT	
rs1179387	9182	AAGGTAAGCAGAGAT ACGTAAATTAT		GGTTATCTTTGGGTAGT AGGATTATA	

[00153] A typical PCR reaction included 14.24 μl of water, 2.23 μl of PCR buffer, 1.38 μl of 1.5 mM $MgCl_2$, 1.12 μl of 0.125 mM dNTPs, 0.45 μl of the forward primer at a 0.2 μM concentration, 0.45 μl of the reverse primer at a 0.2 μM concentration, 0.13 μl of Taq polymerase (0.003 U/μl), and 2.3 μl of DNA sample at a 0.2 ng/μl concentration, for a total volume of 22.3 μl. The PCR reaction was normally carried out using one step at 95°C for 10 minutes; 45 cycles at 95°C for 30 seconds, 60°C for 45 seconds, and 72°C for 45 seconds; one step at 72°C for 5 minutes; and then finalizing the reaction at 22°C.

[00154] After the PCR reaction was completed, an extension oligonucleotide was hybridized to the PCR product. Extension oligonucleotides are reported in Table 3.

Table 3

Position in SEQ ID NO:1	Extension Oligonucleotide
4050	TGAGATGGGAGGATCT (antisense)
4689	ACTGGGAACCTCGA (antisense)
6282	GCTGATGCCGCTG (antisense)
6358	GGAGTGACCCCTT
7256	ACACATGACAACTGCTA
7300	GGTGTGGGTGTACGG (antisense)
7301	GGTGTGGGTGTACGG (antisense)

Position in SEQ ID NO:1	Extension Oligonucleotide
7328	CCACACCTATTCATACTC
8062	CTTAGGCAGGAGAATC (antisense)
9182	GTAATGCAACTTCAAAC

[00155] The extension oligonucleotide was complementary to the amplified target up to but not including the polymorphism (except for examination of polymorphic sites rs2009391 and rs5635, where the extension oligonucleotide terminated one base pair to the polymorphic position), and was enzymatically extended one or a few bases through the polymorphic site. In the extension phase of the assay, a single dNTP was added to the reaction, and pyrophosphate was generated if the dNTP was added to the extension oligonucleotide. ATP sulfurylase present in the reaction mixture utilized the pyrophosphate in conjunction with APS to generate ATP. ATP drove the luciferase-catalyzed conversion of luciferin to oxyluciferin, which generated the release of visible light measured by a CCD camera. A graphic representation was generated showing a peak corresponding to the amount of light emitted, where the light was proportional to the amount of nucleotide incorporated into the extension oligonucleotide. dATP was not added to the reaction, and instead, was replaced by dATPγS, which was not turned over by luciferase. Apyrase was added to the reaction to degrade unincorporated dNTP and ATP sulfurylase-generated ATP, and when the apyrase reaction was complete, another dNTP was optionally added to the reaction for another extension phase.

[00156] An alternative assay involved a MassARRAY™ system (Sequenom, Inc.), which was utilized to perform SNP genotyping in a high-throughput fashion. This genotyping platform was complemented by a homogeneous, single-tube assay method (hME™ or homogeneous MassEXTEND™ (Sequenom, Inc.)) in which two genotyping primers anneal to and amplify a genomic target surrounding a polymorphic site of interest. A third primer (the MassEXTEND™ primer), which is complementary to the amplified target up to but not including the polymorphism, was then enzymatically extended one or a few bases through the polymorphic site and then terminated.

[00157] For each polymorphism, SpectroDESIGNER™ software (Sequenom, Inc.) was used to generate a set of PCR primers and a MassEXTEND™ primer was used to genotype the polymorphism. Table 4 shows PCR primers and Table 5 shows extension primers used for analyzing polymorphisms. The initial PCR amplification reaction was performed in a 5 μl total volume containing 1X PCR buffer with 1.5 mM $MgCl_2$ (Qiagen), 200 μM each of dATP, dGTP, dCTP, dTTP (Gibco-BRL), 2.5 ng of genomic DNA, 0.1 units of HotStar DNA polymerase (Qiagen), and 200 nM each of forward and reverse PCR primers specific for the polymorphic region of interest.

Table 4

Reference SNP ID	Position in SEQ ID NO:1	First PCR primer	SEQ ID NO:	Second PCR primer	SEQ ID NO:
rs2701632	436	ACCCACTTAGCATCCT TCAG		TCTTATGTGGGTTCC TTGGG	
rs2701631	839	TGTGGCCATTGTGACT GAGA		GCCCGGGTGACAGA GTG	
rs5633	6358	TGTGGCAGTTCCGCAA AATG		AGTAGCAGCCGTAGT TGTTG	
rs2070873	6653	ACCCCGTTAGAGATGG AAAC		CTGTTGCTACATTCT GCCAC	
rs5637	7328	AATTTCTGCTGGACAA CCCG		CCTACTGCTACAGGT GATTG	
rs1179387	9182	CAAGCCAAAAGTAATG CAAC		GGATTATAGATGCCT TCCAC	
rs2066539	10164	TCATCTCACACTGTAC TCTC		CAATATCCAAACATG AGGTC	
rs2701629	11649	GACAGAGAGAGACAC TATCT		GAAATGCAAGCTGTT ATTGG	

[00158] Samples were incubated at 95°C for 15 minutes, followed by 45 cycles of 95°C for 20 seconds, 56°C for 30 seconds, and 72°C for 1 minute, finishing with a 3 minute final extension at 72°C. Following amplification, shrimp alkaline phosphatase (SAP) (0.3 units in a 2 μl volume) (Amersham Pharmacia) was added to each reaction (total reaction volume was 7 μl) to remove any residual dNTPs that were not consumed in the PCR step. Samples were incubated for 20 minutes at 37°C, followed by 5 minutes at 85°C to denature the SAP.

[00159] Once the SAP reaction was complete, a primer extension reaction was initiated by adding a polymorphism-specific MassEXTEND™ primer cocktail to each sample. Each MassEXTEND™ cocktail included a specific combination of dideoxynucleotides (ddNTPs) and deoxynucleotides (dNTPs) used to distinguish polymorphic alleles from one another. In Table 5, ddNTPs are shown and the fourth nucleotide not shown is the dNTP (*e.g.*, in the first row A, C and G are ddNTPs and T is the dNTP).

Table 5

Position in SEQ ID NO:1	Extend Probe	SEQ ID NO:	Termination Mix
436	TTAGCATCCTTCAGGCCTAAA		A,C,G
839	GACTCTGCCTCAAAATAAATAAAA (antisense)		C,G,T
6358	GCCGTAGTTGTTGTATTCCAA (antisense)		A,C,T
6653	GTGCAAAACAGTGGGCGATGCT		A,C,T
7328	TGATTGCCGAGCCAGAGCA (antisense)		A,C,G
9182	TTTCCATAATAGATATTTATGTAG (antisense)		C,G,T

Position in SEQ ID NO:1	Extend Probe	SEQ ID NO:	Termination Mix
10164	CACTGTACTCTCCAATAAAGCACC		A,C,G
11649	CAAACAAACACACACACAAAAC		C,G,T

[00160] The MassEXTEND™ reaction was performed in a total volume of 9 μl, with the addition of 1X ThermoSequenase buffer, 0.576 units of ThermoSequenase (Amersham Pharmacia), 600 nM MassEXTEND™ primer, 2 mM of ddATP and/or ddCTP and/or ddGTP and/or ddTTP, and 2 mM of dATP or dCTP or dGTP or dTTP. The deoxy nucleotide (dNTP) used in the assay normally was complementary to the nucleotide at the polymorphic site in the amplicon. Samples were incubated at 94°C for 2 minutes, followed by 55 cycles of 5 seconds at 94°C, 5 seconds at 52°C, and 5 seconds at 72°C.

[00161] Following incubation, samples were desalted by adding 16 μl of water (total reaction volume was 25 μl), 3 mg of SpectroCLEAN™ sample cleaning beads (Sequenom, Inc.) and allowed to incubate for 3 minutes with rotation. Samples were then robotically dispensed using a piezoelectric dispensing device (SpectroJET™ (Sequenom, Inc.)) onto either 96-spot or 384-spot silicon chips containing a matrix that crystallized each sample (SpectroCHIP™ (Sequenom, Inc.)). Subsequently, MALDI-TOF mass spectrometry (Biflex and Autoflex MALDI-TOF mass spectrometers (Bruker Daltonics) can be used) and SpectroTYPER RT™ software (Sequenom, Inc.) were used to analyze and interpret the SNP genotype for each sample.

SNP Verification

[00162] Polymorphisms identified in the publicly available database were verified by detecting the presence or absence of each polymorphism across six individuals from Sweden (including PCR negative control and one sequence primer extension control). Where a polymorphism was present in two or more of the individuals, the polymorphism was designated as a statistically significant SNP and genotyped across the test population. Where the polymorphism was not identified in any of the six individuals, it was further examined in a population of thirty Caucasian blood donors from Sweden. In this group of thirty individuals, a polymorphism having a frequency of 10% or greater was designated as a statistically significant SNP and genotyped across the test population. The probability of not identifying a minor allele variant represented in 10% or more of a population was calculated as being about 0.2% when samples from 30 individuals are analyzed, where it was estimated that 19% of individuals in the total population would be carriers for the minor allele assuming a large population and no selection pressure. Also, polymorphisms were verified in a group of samples isolated from 92 individuals originating from the state of Utah in the United States, Venezuela and France (Coriell cell repositories).

[00163] The following polymorphisms reported in the dbSNP database were identified as being polymorphic (*i.e.*, statistically significant) in the verification studies: rs2701632, rs200931, rs5631, rs5632, rs5634, rs5637, rs1186217, rs1179387, rs2701629, and rs2070873. Polymorphisms reported in the dbSNP database as rs2701631, rs2066539, rs5633, rs5635 and rs5636 were identified as not polymorphic when tested in seventeen individuals.

Genotype Analysis

[00164] Among the verified SNPs, Table 6 depicts two SNPs that were strongly associated with reduced fat deposition. Allele frequency is noted in the second column and the allele indicated in bold type is the allele associated with decreased central fat deposition. These positions were found to be in strong linkage disequilibrium (LD). Statistical significance of each association was determined by the Monks-Kaplan test using a point-wise analysis (Monks & Kaplan, *Am. J. Hum. Genet. 66*: 576-592 (2000)).

Table 6

SNP Position in SEQ ID NO: 1	Allele Frequency	Statistical Significance using Monk-Kaplan Analysis
7328	**A = 0.15588** G = 0.84412	p = 0.00669
9182	**G = 0.14776** T = 0.85224	p = 0.00688

[00165] Correction for multiple testing was also carried out for PLA2G1B, after removal of the other SNPs from the dataset. The value obtained for multiple correction in this manner was p=0.0859.

[00166] Haplotype analysis was performed using a program known as QPDT (Martin *et al., Amer. J. Human Genetics, 67*: 146-54 (2000)), which utilizes the EM algorithm (Dempster *et al., J. Royal Statistical Soc., B39*: 1-38 (1977)). The program was utilized to assign haplotypes based on likelihood of maximization. Table 7 shows possible haplotypes for four SNPs in the PLA2G1B gene and estimated frequencies for each.

Table 7

Allele	Nucleotide Position in SEQ ID NO: 1				Frequency
	4050	7256	7328	9182	
H1	G	T	G	T	0.51297
H2	T	T	G	T	0.29625
H3	T	T	A	G	0.14467
H4	G	C	G	T	0.04292
H5	G	T	A	G	0.00108

[00167] Haplotype versus single position association analysis for the PLA2G1B gene suggested that the H3 haplotype and H5 haplotype were most significantly associated with leanness. These haplotypes are characterized by an A at position 7328 and a G at position 9182.

Example 3

NIDDM Sample Selection

Pooling Strategies

[00168] Samples were placed into one of four groups based on disease status. The four groups were female case samples, female control samples, male case samples, and male control samples. A select set of samples from each group were utilized to generate pools, and one pool was created for each group. Each individual sample in a pool was represented by an equal amount of genomic DNA. For example, where 25 ng of genomic DNA was utilized in each PCR reaction and there were 200 individuals in each pool, each individual would provide 125 pg of genomic DNA. Inclusion or exclusion of samples for a pool was based upon the following criteria and detailed in the tables below. Selection criteria for the study described herein included patient ethnicity and diagnosis with NIDDM. Other phenotypic data collected included GAD antibody concentration, HbA1c concentration, body mass (BMI), patient age, date of primary diagnosis, age of individual as of primary diagnosis (See Table 8 below). Cases with elevated GAD antibody titers and low age of diagnosis were excluded to increase the homogeneity of the diabetes sample in terms of underlying pathogenesis. Controls with elevated HbA1c were excluded to remove any undiagnosed diabetics. Control samples were derived from non-diabetic individuals with no family history of NIDDM. Secondary phenotypes were also measured in the diabetic cases, phenotypes such as HDL , LDL, triglycerides, insulin, C-peptide, nephropathy status, neuropathy status, to name a few, which will allow secondary analysis of the cases the be performed in order to elucidate the potential pathway of the disease gene.

Table 8

Exclusion Criteria	No. of individuals fulfilling exclusion criteria	Actual no. of samples excluded after each stage	No. of samples remaining
ALL SAMPLES Lack of availability of sample	34	34	1591
ALL SAMPLES Non-German ethnicity	261	239	1352
CASES GAD Ab >0.9	102	84	1268
CONTROLS HbA1c ≥ 6 or BMI > 40	21	20	1248
CASES age < 90	17	6	1242
CASES Age of Diagnosis < 35, **CONTROLS** Family History of Diabetes	150 170	203	1039
CONTROLS Age-matching to case pool	43	43	996

[00169] The selection process yielded the pools set forth in Table 9, which were used in the studies described herein.

Table 9

	Female case	**Female control**	**Male cases**	**Male control**
Pool size (Number)	244	244	254	254
Pool Criteria (ex: case/control)	case	control	case	control
Mean Age (ex: years)	52.49	49.02	49.78	50.57

Example 4

Association of Polymorphic Variations with NIDDM

[00170] Blood samples were taken from individuals in the study population described in Example 3. Genomic DNA was extracted from these blood samples using standard techniques (BACC2 DNA extraction kit (Nucleon Biosciences)) and subjected to analysis. Based upon the the coexistence of all of the following or differing combinations of central fat, hypertension, glucose intolerance, dyslipidemia (elevated triglycerides and low HDL cholesterol), and impaired insulin stimulated glucose uptake ("insulin resistance") in a common disorder referred to as syndrome X, it was postulated that

polymorphic variants associated with the development of cental obesity would also be associated with NIDDM.

[00171] The SNP at position 7256 of SEQ ID NO: 1 was also allelotyped and genotyped in NIDDM and non-NIDDM patients from the pool described above (see Example 4). The following PCR primers were used: ACGTTGGATGGGGTTGTCCAGCAGAAATTTAC (forward PCR primer) and ACGTTGGATGCTTTCCAGGTGCTGCCAG (reverse PCR primer); and AGACACATGACAACTGCTA (extend primer).

Genotype Analysis

[00172] The SNP at position 7256 of SEQ ID NO: 1 was allelotyped and genotyped in NIDDM and non-NIDDM patients as described in Example 2. Table 10 shows the allelotyping results for the SNP at position 7256. Allele frequency is noted in the second column and the allele indicated in bold type is the allele associated with NIDDM. Table 11 shows the genotyping results for the SNP at position 7256. Genotype frequency in cases and controls is noted in columns 2, 3 and 4. Statistical significance of each association was determined by the Pearson Chi-squared test. .

Table 10

SNP Position 7256 in SEQ ID NO: 1	**Allele Frequency Cases**	**Allele Frequency Controls**	**Statistical Significance**
Females	T = 0.924 C = 0.076	T = 0.934 C = 0.066	p = 0.736
Males	T = 0.895 C = 0.105	T = 0.946 C = 0.054	p = 0.048

Table 11

SNP Position 7256 in SEQ ID NO: 1	TT	TC	CC	Statistical Significance
Case Female	0.886	0.114	0.000	p = 0.461
Control Female	0.901	0.094	0.005	
Case Male	0.915	0.077	0.008	p = 0.022
Control Male	0.855	0.145	0.000	

[00173] Both allelotyping and genotyping analysis revealed that a cytosine at position 7256 of SEQ ID NO: 1 is associated with NIDDM (most significantly in males). Interestingly, a guanine at position 7328 and a thymine at position 9182 of SEQ ID NO: 1 were found to be associated with central obesity (see Example 2). Therefore, the data demonstrates these SNP serve as a marker for an increased risk of developing obesity or diabetes either separately or together as part of a greater metabolic syndrome.

Example 5

PLA2G1B Tissue Expression Profiles

[00174] PLA2G1B expression levels were determined in tissues of Israeli sand rats (*Psamommys obesus*) by detecting RNA transcribed from the PLA2G1B gene. *P. obesus* is a polygenic animal model ideal for the study of obesity and type 2 diabetes. *P. obesus* displays a range of pathophysiologic phenotypic responses when fed a standard laboratory diet *ad libitum* and animals were classified into four groups as set forth in Table 12.

Table 12

Group	Phenotype	Plasma glucose/Insulin
Group A	Healthy	Normoglycemic/ normoinsulinemic
Group B	Insulin resistant	Normoglycemic/ hyperinsulinemic
Group C	Diabetic/Obese	Hyperglycemic/ hyperinsulinemic
Group D	Diabetic/Obese	Hyperglycemic/ hypoinsulinemic

[00175] Studies were typically performed on group A, B and C animals as group D animals developed decompensated diabetes when their pancreas failed, leading to rapid death. Animals were classified at 16 weeks age following body weight, blood glucose and plasma insulin measurements at 8, 12 and 16 weeks. Body weight, blood glucose, and blood insulin were measured in grams, mmol/L, and

mμ/L, respectively. Animals were considered lean at 12 weeks if their body weight was less than 180 grams and obese when body weight was greater than 200 grams. Animals were considered normoglycemic and normoinsulinemic if their blood glucose levels were less than 8.0 mmol/L and insulin levels were less than 150 mμ/L. Animals were classified as hyperinsulinemic if their blood insulin levels were equal to or greater than 150 mμ/L. Animals were further classified as diabetic if their blood glucose levels were equal to or greater than 8 mmol/L.

[00176] PLA2G1B tissue distribution expression profiles were studied in male *P. obesus* group A animals (lean and healthy) and the results are depicted in Figures 3A-3D. Animals were normally fasted for two hours prior to tissue harvesting. As shown in Figures 3A-3D, PLA2G1B expression was highest in stomach tissue, and expressed at lower levels in pancreatic, lung, and adrenal tissue. Expression was also observed in the large and small intestine.

[00177] Metabolically-linked tissues, such as liver, fat pads, skeletal muscle, hypothalamus, pancreas, and stomach tissues, were targeted for analysis of differential gene expression of PLA2G1B following normal feeding or overnight fasting conditions. In addition, data relating to blood glucose, plasma insulin, body weight, and body fat from the animals were correlated against gene expression using t-test analysis. From these studies, it was determined that PLA2G1B expression in the hypothalamus was significantly greater in group C fasted animals as compared to group A fasted animals ($p=0.033$) and group B fasted animals ($p=0.02$) using a parametric t-test. (See Figure 4A). Also, hypothalamus PLA2G1B expression in group A animals that were fed normally was greater than in fasted group A animals ($p=0.052$). (See Figure 4B). In addition, hypothalamus PLA2G1B expression in fasted animals was positively associated with body weight ($p=0.028$) and plasma insulin levels ($p=0.014$) using a parametric Pearson comparison. (See Figures 4C and 4D).

[00178] In the liver, PLA2G1B expression in group A fasted animals tended to be lower than group B fasted animals ($p=0.072$) and group C fasted animals ($p=0.023$) using a Games-Howell parametric method of multiple comparisons. (See Figure 4E). In addition, liver PLA2G1B expression in normally fed group A animals were lower than normally fed group C animals ($p=0.067$). (See Figure 4F) Also, there were positive associations between liver PLA2G1B expression in fasted animals with body weight ($p = 0.005$), blood insulin ($p = 0.013$ both parametric correlations), and blood glucose ($p=0.023$, nonparametric correlation). (See Figures 4G, 4H and 4I). Further, there was a positive association between liver PLA2G1B expression in fed animals with body weight ($p=0.013$). (See Figure 4J).

[00179] In the pancreas, a significant difference in PLA2G1B expression was observed between control and energy-restricted groups ($p = 0.036$, t-test). (See Figure 4K). There were no correlations between body weight, blood glucose, and blood insulin with pancreatic PLA2G1B expression.

[00180] In subscapular fat, PLA2G1B expression in the fasted animals was significantly greater than normally fed animals ($p = 0.038$ t-test). (See Figure 4L). In red gastrocnemius muscle, kidney, and stomach tissues, however, there were no significant differences in PLA2G1B expression between fed groups and between fasted groups, and no correlations between body weight, blood glucose, and blood insulin with PLA2G1B expression.

[00181] Gene expression was quantified using a TaqMan™ PCR system (ABI PrismTM 7700 Sequence Detection System, Perkin-Elmer Applied Biosystems, Norwalk, USA) and was determined relative to an endogenous control gene, cyclophilin. cDNA was synthesized by subjecting one microgram of total RNA to a reverse transcription reaction using SuperScript II RNase H- Reverse Transcriptase (Invitrogen) according to manufacturer's instructions (*see* http address www.invitrogen.com/Content/ World/11904018.pdf). In this reverse transcriptase PCR (RT-PCR) procedure, the following contents were added to a nuclease-free microcentrifuge tube: 1 µl Oligo (dT)12-18 (500 µg/ml); 1 µg total RNA; 1 µl 10 mM dNTP mix (10 mM each dATP, dGTP, dCTP and dTTP at neutral pH); sterile, distilled water to 12 µl. The mixture was heated to 65°C for 5 minutes and quickly chilled on ice for at least 2 minutes.

[00182] Contents of the tube were collected by brief centrifugation and the following were added to complete a 20-µl reaction volume: 2 µl 10X First-Strand Buffer; 4 µl 25 mM $MgCl_2$; 2 µl 0.1 M DTT; 1 µl RNaseOUT Recombinant Ribonuclease Inhibitor (40 units/µl); 1 µl (200 units) of and SUPERSCRIPT II. The mixture was incubated at 45°C for 50 minutes and then the reaction was inactivated by heating at 70°C for 15 minutes. To remove RNA complementary to the cDNA, 1 µl (2 units) of *E. coli* RNase H was added and incubated at 37°C for 20 minutes. The resulting mixtures were transferred to 0.5 ml tubes and stored at -20°C.

[00183] Oligonucleotide primers were designed based upon the *P. obesus* sequence using Primer Express software (version 1.5), which was obtained at the http address docs.appliedbiosystems.com/pebiodocs/04303014.pdf. For PCR reactions, forward primers having the sequences GCTGTGTGGCAGTTCCGCAA; GTTCCGCAATATGATCAAGTGC; GATGAAACTCCTTCTGCTGGCTG; and SAAGATGAAACTCCTTCTGCTG were utilized in conjunction with reverse primers having the sequences GGTGAAATAAGACAGCAAGG; GGAGAANCAGATGGCGGCCT; CGGTCACAGTTGCAGATGAAG; GGAAGTGGGGTGACAGCCTAACA; and GGTGACAGSCTAACAGWNTTTC, where S is G or C; N is C, G, T, or A; and W is A or T. Also, another forward primer having the sequence 5'-GCACCCCAGTGGACGAATT-3' and a reverse primer having the sequence 5'-TCAGCCTCTTGGCCTTAGTGTAG-3' yielded an amplicon that was 70 base pairs in length and were used for RT-PCR. Primers for the endogenous control gene, cyclophilin, were designed based on the *P.*

obesus sequence. Primer sequence specificity was confirmed by comparing the primer sequences against the GenBank nucleotide sequence for PLA2G1B using BLAST. Primers were synthesized at a 40 nmole concentration and purified by using a reverse-phase cartridge (GeneWorks, Australia).

[00184] The ability of the primers to operate in a quantitative PCR process was next determined. A standard curve was generated based upon threshold cycles (Ct = threshold cycle) for serially diluted samples. cDNA was serially diluted from a 1:2 dilution to a 1:16 dilution, and the standard curve included an undiluted sample and a "no template control" (contains no cDNA). A standard curve was also generated using primers specific for the endogenous control gene (cyclophilin). These samples are set-up in duplicate using the following: 12.5 µl of SYBR Green Universal PCR master mix (cat# 4304437, http address docs.appliedbiosystems.com/pebiodocs/00777601.pdf); 2.5 µl of forward primer (1µM, diluted in nuclease-free water); 2.5 µl of reverse primer (1µM, diluted in nuclease-free water); 2.0 µl of cDNA (neat or diluted); and 5.5 µl of water (nuclease-free) for a total volume of 25µl.

[00185] The PCR program recommended for the ABI Prism 7700 procedure was utilized and the baseline was calculated based upon cycles 3 to 15 and the amplification plot was based upon cycles 16 to 40. A threshold level was set following examination of a semi-log view of the plot. The Ct values for each duplicate were examined to ensure they did not differ by more than one Ct unit. The Ct values were eliminated or the experiment was repeated if they differed by more than one Ct unit. Samples were run on an agarose gel to identify product formation and whether or not primer-dimers or non-specific priming occurred. While the primer concentration could have been optimized if required, it was determined that 100 nM of each primer (final concentration) was adequate.

[00186] Following the primer efficiency determination, a real time PCR run was executed. The conditions utilized were as described above except that cDNAs were diluted 1:8 and products were not confirmed on an agarose gel. Final values were then calculated using the relation $2^{-\Delta Ct}$, where ΔCt is Ct of cyclophilin subtracted from Ct of the gene of interest. Gene expression values were calculated as arbitrary units, and Ct values for cyclophilin in treated samples (*e.g.*, in fasted tissues) were further examined to determine whether endogenous control of gene expression was altered. This analysis yielded quantified and standardized gene expression values for the amount of cDNA in each reaction.

Example 6

In Vitro Production of PLA2G1B Polypeptides

[00187] PLA2G1B cDNA is cloned into a pIVEX 2.3-MCS vector (Roche Biochem) using a directional cloning method. A PLA2G1B cDNA insert is prepared using PCR with forward and reverse primers having 5' restriction site tags (in frame) and 5-6 additional nucleotides in addition to 3' gene-specific portions, the latter of which is typically about twenty to about twenty-five base pairs in length.

A Sal I restriction site is introduced by the forward primer and a Sma I restriction site is introduced by the reverse primer. The ends of PLA2G1B PCR products are cut with the corresponding restriction enzymes (*i.e.*, Sal I and Sma I) and the products are gel-purified. The pIVEX 2.3-MCS vector is linearized using the same restriction enzymes, and the fragment with the correct sized fragment is isolated by gel-purification. Purified PLA2G1B PCR product is ligated into the linearized pIVEX 2.3-MCS vector and *E. coli* cells transformed for plasmid amplification. The newly constructed expression vector is verified by restriction mapping and used for protein production.

[00188] *E. coli* lysate is reconstituted with 0.25 ml of Reconstitution Buffer, the Reaction Mix is reconstituted with 0.8 ml of Reconstitution Buffer; the Feeding Mix is reconstituted with 10.5 ml of Reconstitution Buffer; and the Energy Mix is reconstituted with 0.6 ml of Reconstitution Buffer. 0.5 ml of the Energy Mix was added to the Feeding Mix to obtain the Feeding Solution. 0.75 ml of Reaction Mix, 50 µl of Energy Mix, and 10 µg of the PLA2G1B template DNA is added to the *E. coli* lysate.

[00189] Using the reaction device (Roche Biochem), 1 ml of the Reaction Solution is loaded into the reaction compartment. The reaction device is turned upside-down and 10 ml of the Feeding Solution is loaded into the feeding compartment. All lids are closed and the reaction device is loaded into the RTS500 instrument. The instrument is run at 30°C for 24 hours with a stir bar speed of 150 rpm. The pIVEX 2.3 MCS vector includes a nucleotide sequence that encodes six consecutive histidine amino acids on the C-terminal end of the PLA2G1B polypeptide for the purpose of protein purification. PLA2G1B polypeptide is purified by contacting the contents of reaction device with resin modified with Ni^{2+} ions. PLA2G1B polypeptide is eluted from the resin with a solution containing free Ni^{2+} ions.

Example 7
Cellular Production of PLA2G1B Polypeptides

[00190] PLA2G1B nucleic acids are cloned into DNA plasmids having phage recombination cites and PLA2G1B polypeptides and polypeptide variants are expressed therefrom in a variety of host cells. alpha phage genomic DNA contains short sequences known as attP sites, and *E. coli* genomic DNA contains unique, short sequences known as attB sites. These regions share homology, allowing for integration of phage DNA into *E. coli* via directional, site-specific recombination using the phage protein Int and the *E. coli* protein IHF. Integration produces two new att sites, L and R, which flank the inserted prophage DNA. Phage excision from *E. coli* genomic DNA can also be accomplished using these two proteins with the addition of a second phage protein, Xis. DNA vectors have been produced where the integration/excision process is modified to allow for the directional integration or excision of a target DNA fragment into a backbone vector in a rapid *in vitro* reaction (Gateway™ Technology (Invitrogen, Inc.)).

[00191] A first step is to transfer the PLA2G1B nucleic acid insert into a shuttle vector that contains attL sites surrounding the negative selection gene, ccdB (*e.g.* pENTER vector, Invitrogen, Inc.). This transfer process is accomplished by digesting the PLA2G1B nucleic acid from a DNA vector used for sequencing, and to ligate it into the multicloning site of the shuttle vector, which will place it between the two attL sites while removing the negative selection gene ccdB. A second method is to amplify the PLA2G1B nucleic acid by the polymerase chain reaction (PCR) with primers containing attB sites. The amplified fragment then is integrated into the shuttle vector using Int and IHF. A third method is to utilize a topoisomerase-mediated process, in which the PLA2G1B nucleic acid is amplified via PCR using gene-specific primers with the 5' upstream primer containing an additional CACC sequence (*e.g.*, TOPO® expression kit (Invitrogen, Inc.)). In conjunction with Topoisomerase I, the PCR amplified fragment can be cloned into the shuttle vector via the attL sites in the correct orientation.

[00192] Once the PLA2G1B nucleic acid is transferred into the shuttle vector, it can be cloned into an expression vector having attR sites. Several vectors containing attR sites for expression of PLA2G1B polypeptide as a native polypeptide, N-fusion polypeptide, and C-fusion polypeptides are commercially available (*e.g.*, pDEST (Invitrogen, Inc.)), and any vector can be converted into an expression vector for receiving a PLA2G1B nucleic acid from the shuttle vector by introducing an insert having an attR site flanked by an antibiotic resistant gene for selection using the standard methods described above. Transfer of the PLA2G1B nucleic acid from the shuttle vector is accomplished by directional recombination using Int, IHF, and Xis (LR clonase). Then the desired sequence can be transferred to an expression vector by carrying out a one hour incubation at room temperature with Int, IHF, and Xis, a ten minute incubation at 37°C with proteinase K, transforming bacteria and allowing expression for one hour, and then plating on selective media. Generally, 90% cloning efficiency is achieved by this method. Examples of expression vectors are pDEST 14 bacterial expression vector with att7 promoter, pDEST 15 bacterial expression vector with a T7 promoter and a N-terminal GST tag, pDEST 17 bacterial vector with a T7 promoter and a N-terminal polyhistidine affinity tag, and pDEST 12.2 mammalian expression vector with a CMV promoter and neo resistance gene. These expression vectors or others like them are transformed or transfected into cells for expression of the PLA2G1B polypeptide or polypeptide variants. These expression vectors are often transfected, for example, into murine-transformed a adipocyte cell line 3T3-L1, (ATCC), human embryonic kidney cell line 293, and rat cardiomyocyte cell line H9C2.

Example 8

Cellular Assay for Screening PLA2G1B Interacting Fat Reduction Drug Candidates

[00193] General PLA2 assay strategies are known (Reynolds *et al., Methods in Enzymology 197*: 3-23 (1991)). Sensitive and practical assays include radioactive and spectrophotometric assays. An assay optionally employing chromogenic or spectrometric detection (Yu *et al., Methods in Enzymology* 197: 65-75 (1991)) is often utilized for determining whether test molecules interact with PLA2G1B in a high throughput format, typically with inclusion of bile acids or other anionic detergents. The assay format has been modified with minor variations to assay the non-pancreatic GIIA PLA2 from human synovial fluid in a high throughput format (Reynolds *et al., Analytical Biochemistry 204*:190-197 (1992)).

[00194] A similar spectrophotometric assay was developed for GIVA PLA2 (Reynolds *et al. Anal. Biochem. 217*:25-32 (1994)) and is utilized to determine whether a test molecule interacts with PLA2G1B. This assay is often utilized in conjunction with a microtitre plate and plate reader in a high throughput format. In the assay, PLA2 function is monitored using a ThioPC/Triton X-100 substrate solution. An appropriate volume of ThioPC in chloroform solution is evaporated to dryness under a stream of N_2. Triton X-100 (8 mM) in 2X assay buffer (160 mM HEPES, pH 7.4, 300 mM NaCl, 20 mM $CaCl_2$, 2 mg/ml BSA) is added to the dried lipid in one-half the desired final volume to give a 2-fold concentrated substrate solution. This solution is bath-sonicated for 1 minute to loosen dried ThioPC from the walls of the vial and then probe-sonicated on ice (20 seconds on ice, 20 seconds off ice) for 3 minutes. The solution is then warmed to 40°C and warmed glycerol equivalent to 30% of the final volume was added. The solution is then brought to the desired final volume with deionized H_2O. The final assay contains 2 mM ThioPC, 4 mM Triton® X-100 and 30% glycerol in 80 mM HEPES, pH 7.4, 150 mM NaCl, 10 mM $CaCl_2$ and 1 mg/ml BSA.

[00195] The substrate is then aliquotted, in 200 µl increments, into the wells of a 96-well plate and equilibrated for 5 minutes at 37°C To initiate the reaction, 500 ng PLA2 (purified, recombinant human), in a 5 µl volume of 1X assay buffer, is added to the wells, the plate is shaken 20 seconds on high to mix, and then incubated for 60 minutes at 37°C. For controls, buffer rather than enzyme was added to some wells.

[00196] After 60 minutes, 10 µl of a 25 mM DTNB/475 mM EGTA mixture is added to all substrate containing wells to quench the reaction and initiate the color development. The DTNB/EGTA mixture is prepared just prior to use by combining equal volumes of 50 mM DTNB in 0.5M Tris, pH 7.4 and 950 mM EGTA, pH 7.2. After adding the DTNB/EGTA, the plate is once again shaken 20 seconds on high and allowed to incubate for an additional 3 minutes to give the DTNB chromophore time to fully develop prior to reading the plate. The absorbance is measured using a dual wavelength option (405 nm to 620 nm) to correct for light scattering. The results obtained with this dual wavelength option are similar to

those obtained using a single wavelength (405 nm) but are more reproducible. The average absorbance of the controls is subtracted from that of the enzyme-containing wells to correct for the absorbance due to the substrate, DTNB, and EGTA. The difference in absorbance is used to calculate enzyme activity. The data was reported ± the standard deviation. Specific activity is calculated using ε_{405} for DTNB of 12,800 M^{cm-1} and a path length of 0.47 cm for a 215 µl final total volume. The path length in these plates is dependent on the assay volume and was calculated by measuring the absorbance of several concentrations of bromothymol blue, where the path length equals the absorbance observed on the plate reader divided by the absorbance observed for the same solution in the spectrophotometer in 1 cm cuvettes. A short burst of activity is often observed in the firsts 5 minutes followed by a more linear phase from 5 to 60 minutes. Further details concerning this assay are disclosed in U.S. Patent No. 5,464,754. This assay also can be carried out using a modified phosphocholine substrate as is used when assaying cobra venom PLA2 molecules.

[00197] Further, assays described in (Yang *et al., J. Neurochemistry 73*:1278-1287 (1999)) readily can be applied to distinguish secreted PLA2 molecules (*e.g.*, PLA2G1B) in tissues from other PLA2 forms.

Example 9

In Vivo Assay for Screening Fat Reduction Drug Candidates

[00198] Test molecules are screened for fat reduction activity by administering molecules which interact with PLA2G1B to Israeli sand rats (*P. obsesus*), which is an accepted *in vivo* model for obesity, and observing the effect of the molecule on such parameters as weight, dimensions, and/or fat content. Molecules may be administered to obese animals and/or non-obese animals. These animals are grouped into four sets (Table 8), where group D animals have high morbidity and are not typically used in studies.

[00199] The Israeli sand rat is maintained on an *ad libitum* diet of a standard lab chow that is high in energy. This polygenic animal displays in response to this diet a range of body weights, plasma insulin and blood glucose levels. Normally, eight controlled animals and eight treated animals are included for groups A, B and C, giving a total of 48 animals for each study.

[00200] The test molecule is delivered to the animals by intraperitoneal injection; intravenous injection; intragastrical administration, in which case twice as many animals per group should be used since the method of administration is more stressful and leads to a higher motility rate; continuous infusion using an osmotic pump; and orally *ad libitum*, which is the least stressful as the test molecule is added to food and the amount of consumed is measured. Often DMSO or water is used as a vehicle accompanying the test molecule and 10 µg to 1000 µg of test molecule per kilogram of the animal is typically administered.

[00201] The length of the study is typically one to seven days. During the study, several parameters are measured, including body weight (daily measurements); food intake (daily measurements); blood glucose levels (before and after the study); plasma insulin levels (before and after the study); circulating blood metabolites such as leptin, cortisol, triglycerides and free fatty acids (before and after the study); percent body fat (weighing fat pads at the end of the study); quantification of gene expression in tissues such as the pancreas, mesenteric fat, stomach and small intestine (at the end of the study); and measurements of PLA2G1B activity in tissues such as pancreas, mesenteric fact, stomach, and small intestine using methods described in Example 8 (before and/or after the study). Animals are sacrificed by anaesthetic overdose and tissues are harvested and rapidly frozen. RNA is extracted from half of each harvested tissue and PLA2G1B polypeptide extracts are sometimes generated from the other half.

[00202] Modifications may be made to the foregoing without departing from the basic aspects of the invention. Although the invention has been described in substantial detail with reference to one or more specific embodiments, those of skill in the art will recognize that changes may be made to the embodiments specifically disclosed in this application, yet these modifications and improvements are within the scope and spirit of the invention, as set forth in the claims which follow. All publications or patent documents cited in this specification are incorporated herein by reference as if each such publication or document was specifically and individually indicated to be incorporated herein by reference.

[00203] Citation of the above publications or documents is not intended as an admission that any of the foregoing is pertinent prior art, nor does it constitute any admission as to the contents or date of these publications or documents. U.S. patents and other publications referenced herein are hereby incorporated by reference.